WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/134215

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

SOME FOUNDATIONAL PROBLEMS in the THEORY of MEASUREMENT

by

R. J. Horton, MA, BSc.

A thesis presented for the degree of

Doctor of Philosophy

to the

University of Warwick

Department of Philosophy,
University of Warwick.

Department of Applied Physical Sciences,
Coventry (Lanchester) Polytechnic.

October 1986

CONTENTS

Acknowledgements

There are a number of people from whom I have received help in the preparation of this thesis. I would like to thank Sue Haack for a useful discussion and some advice on reading material. I am greatly indebted to Greg Hunt for numerous sessions spent discussing and criticizing my half formed ideas, and for his kindly encouragement throughout the work. Most especially I wish to thank my supervisor David Miller for all his help and generous support. He has been painstaking and unstinting in offering advice and criticism, and I am very conscious of how much poorer the finished product would have been without his guidance. Needless to say the remaining faults are entirely of my own making. In the final year I was greatly helped by being relieved of a substantial part of my teaching in the Polytechnic, and I wish to express my gratitude to colleagues in the Department of Applied Physical Sciences who made this possible. My wife and children helped to produce some of the diagrams and to assemble the final document. But much more than this they provided the moral support without which it is unlikely that the work would have been completed at all.

Summary

The thesis deals with certain aspects of the correspondence between objects and numbers that is exploited in measurement. There are two main areas of investigation.

The first is to do with the construction of scales of measurement, with major emphasis on extensive measurement in physical science. The discussion includes examination of (a) the distinction between those features of scales that are determined by the nature of measurement and those that are determined conventionally, and (b) the role of constraints imposed by characteristics of the objects being measured and constraints imposed by the measurement operations. There is a treatment of experimental error and the limits this sets to our knowledge of the structure of physical systems. The view that extensive scales are designed for *counting* is considered.

The second area relates to the mismatch between empirical relations among physical objects and relations among numbers to which the former are supposed to correspond, a mismatch due to the fact that there are fundamental constraints on the extent to which physical objects can be manipulated for operational purposes. Consequential problems in formal theories of measurement are identified in the thesis, and some new formal structures to accommodate them are proposed.

Material believed to be novel includes (i) some aspects of the treatment of random error (Chapter 2), (ii) some arguments about the conventionality of scales, the relation to counting, and some discussion of counterarguments due to Ellis (Chapter 3), (iii) a suggestion about the classification of scales (Chapter 4), and (iv) the development of certain formal structures, in particular the structures of Definitions 5.4, 6.6. 6.7, 7.3 and 7.13 (Chapters 5, 6 and 7). The development of Definitions 6.6 and 6.7 includes some treatment of the problem of systematic error.

Index of Definitions

CHAPTER ONE

INTRODUCTION

1.0 Preamble

This thesis is concerned with some fundamental aspects of the subject of measurement, especially in the physical sciences. Part is devoted to some problems related to formal theories of measurement.

The accepted view of the foundations of measurement can perhaps be summarized in this way. A property common to some set of objects is quantifiable, and is therefore a suitable candidate for measurement, when a certain kind of correspondence exists between the set of objects and a set of numbers. A key feature of this correspondence is that it involves not simply the objects taken individually but certain kinds of relation among them. That is, a measurable property is seen as a characteristic not of single objects considered in isolation but of their relation to other objects of a similar kind. The property is seen to be quantifiable by virtue of the fact that a set of empirical relations among the objects has an image in a set of corresponding arithmetical relations among the numbers.

Stated as succinctly as possible, my aim is to say something useful about the nature of this correspondence. There are two aspects of interest which together provide the major themes of the thesis. I give an initial brief indication of what these are.

One aspect is the status of the link between objects and numbers. A fundamental issue is whether or not this link is intrinsic to measurable systems. Is it something in the world that is discovered and exploited with the aid of appropriate measuring operations? Or is it instead simply an invention devised to enable us to express empirical information in numerical language? Is the image of the empirical system reflected in the numerical system, or is it simply projected on to it? Field (1980) has argued that the need for numbers in science is entirely a matter of convenience, and that it is possible to give an account of, for example, the structure of space without recourse to numbers at all. Undoubtedly there are features of measurement operations, and of numerical scales, that are settled by convention, and this question may perhaps be understood as one of the demarcation between the conventional and, should there prove to be any, the nonconventional features of given systems of measurement. The first part of the thesis (Chapters 2, 3 and 4) is, generally speaking, concerned with matters related to this question. In particular I wish to examine the idea that a centrally important type of measurement - extensive measurement - can be construed as counting.

The other aspect of the correspondence between objects and numbers is the degree to which numerical relations provide a *faithful* image of the empirical relations. There are important respects in which, either because of fundamental differences in nature between physical objects and numbers, or because of the nature of measurement operations, the correspondence is incomplete. The image, whether reflected or projected, is subject to distortion, and a satisfactory theory of measurement must take this fully into account. The kind of question I am interested in is this. How does the pattern of relations among objects that have some measurable property in common differ from that among the numbers with which they are supposed to correspond, and how does the existence of the differences affect our understanding of the nature of a measurable property? The second part of the thesis (Chapters 5, 6 and 7) is given over largely to questions of this kind.

Questions such as these are of undoubted interest, particularly from the point of view of their significance for the the application of mathematics both in the commonplace examples of everyday life and in the special examples of science and technology. However the amount of philosophical literature on measurement is surprisingly modest. Ellis puts it in the opening sentence of his *Basic Concepts of Measurement*, one of a fairly small number of books in this field, that "measurement is the link between mathematics and science" (1966), and he goes on to comment that, nevertheless, the subject of measurement has, attracted strangely little attention from philosophers of science. Campbell's *Physics: the Elements* (1920) was a pioneer work. This was followed by Bridgman's *Dimensional Analysis* (1931). There appears to have been no other major publications in this genre since Ellis's.

There is on the other hand a considerable body of existing writing devoted to the programme of devising formal stuctures to represent the patterns of relations found in a wide variety of systems of measurement. Much of this makes full use of techniques of abstract algebra and formal languages, and a large body of technical results has been obtained. A comprehensive survey of techniques and applications across the whole range of measurement in both the natural and the social sciences is to be found in *Foundations of Measurement*, Vol. I, by Krantz, Luce, Suppes and Tversky (1971). Though now 15 years old this monumental work is probably still the most important reference source in the field. An important later publication that brings together results of the most recent mathematical developments is Narens, *Abstract Measurement Theory* (1986). As perusal of works such as these quickly reveals, many of the detailed problems that have been encountered have inevitably proved sufficiently interesting and important at a technical level to command attention without regard for wider issues. A great deal of the literature in this area is primarily of mathematical rather than philosophical interest. Nevertheless a prime reason for employing formal techniques is that the process of devising adequate formal structures to represent the material in hand inexorably brings the underlying philosophical issues into sharper focus, and it is in this light that I wish to investigate some specific problems to be found in the formal theory. Before attempting to elaborate on the issues to be dealt with in the body of the thesis I wish to prepare some ground by sketching in the next section what I take to be a standard approach to the construction of a formal theory of measurement. I shall then give, in a subsequent section, a synopsis of the material of later chapters.

1.1 Formal theories of measurement

The general strategy in constructing a formal theory is as follows. The first step is to identify the relevant relations underlying the measurement. The second is to express the essential features of these relations in axiomatic form. The final step is to demonstrate that the axioms are sufficient to support a numerical assignment. I shall elaborate on each of these in the following subsections, 1.1.1 to 1.1.3.

1.1.1 Empirical Relational Structures

In the first stage a suitable *abstract relational structure* is chosen to represent the objects and the empirical relations among them that are associated with the measurable property. A relational structure is a set consisting of a domain A and a (finite) number of relations R_1, R_2,... defined on the domain, and is denoted by:

$$\langle A, R_1, R_2, \ldots, R_n \rangle$$

The empirical relations in question are assumed to be determined by *measurement operations*. It is supposed that for any particular measurable property, attributed to a given set of objects, there is an ideal measurement process that would in principle be sufficient to establish a scale of measurement *ab initio*, and to assign values to the objects according to this scale, without recourse to any pre-existing calibration. In general the process involves one or more *elementary procedures*, specific to the measurable property under consideration, and each of them is associated with a specific relation. A standard example is provided by the case of weight as it is determined by the use of a common balance. For this it is imagined that we have a set of objects whose weights according to any scale are initially unknown, and a balance sufficiently large to accommodate in its pans as many of the objects at one time as we care to place in them. In this example there are two elementary procedures each associated with an empirical relation. Since each of these procedures and their associated relations are representative of types that are central to much of the material of this thesis it is worth paying close attention to them here.

(a) *The Comparison Procedure*

The first procedure is one in which we *compare pairs of objects* - placing them on opposite pans of a balance and observing the outcome - and it is associated with a binary relation, usually denoted in the formal language by $\succsim$. The expression $x \succsim y$ might be interpreted by something like:

> when objects x and y are placed on opposite pans of a balance either the balance remains in equilibrium or else the pan containing object x descends.

[This could be expressed more rigorously. It is a metalinguistic statement and strictly it should be worded "when objects *represented by* x and y are placed...etc.". In the interests of avoiding too tedious a style I shall use the looser form of expression quite freely throughout the text when it leaves the intended meaning sufficiently clear.]

There is just one relational predicate to cover what we might intuitively regard as two separate kinds of outcome, but the advantage of doing this is that a single primitive can then be used to express the result of every test. When needed, the two components of the interpretation can be expressed separately in terms of $\succsim$, and it is convenient to *define* (in terms of $\succsim$) two further terms $\sim$ and $\succ$ in the usual way by:

Def.1.1 (i) $x \sim y$ iff $x \succsim y$ and $y \succsim x$.
(ii) $x \succ y$ iff $x \succsim y$ and not $y \succsim x$.

The interpretations of $\sim$ and $\succ$ are straightforward. $x \sim y$ is:

> when objects x and y are placed in opposite pans of the balance, the balance remains in equilibrium,

and $x \succ y$ is:

> when objects x and y are placed in opposite pans of the balance, the pan containing object x descends.

The importance of $\succsim$ is that in terms of this relation we can order the set of objects according to what we intuitively recognize as weight. If $x \succsim y$ holds then x occupies the same position as, or is higher in the order than y. $\succsim$ is taken to correspond with the arithmetical relation $\geq$ ("is equal to or greater than"). In standard accounts of the foundations of measurement the existence of an ordering procedure associated with a relation of this type is held to be necessary condition for any property to be measurable.

(b) *The Concatenation Procedure*

The second procedure is one in which we *combine objects* - placing two or more together on the same pan of the balance - producing composite objects whose behaviour on the balance may be observed as before. This is associated with a ternary relation. This can be denoted by C where the expression Cxyz might stand for:

> object z is formed by placing object x and object y together in the pan of a balance.

An equivalent more common formulation is in terms of a binary operation, $\circ$, such that Cxyz is replaced by $x \circ y = z$. The concatenation operation, unlike the ordering relation, is not a common feature of all measurable properties but is peculiar to extensive properties, and it is taken to correspond with the arithmetical operation of addition.

Thus on the basis of this analysis we decide on the appropriate structure to represent the property of weight, namely:

$$\langle A, \succsim, C \rangle$$

or, in the more usual formulation:

$$\langle A, \succsim, \circ \rangle$$

Structures such as $\langle A, \succsim \rangle$ and $\langle A, \succsim, \circ \rangle$ that represent systems in which the relations are determined empirically are commonly referred to as *empirical relational structures*.

1.1.2 Definition of a Measurement Structure

The next stage is the formulation of a set of axioms to define the properties of the relational structure. I shall quote here one important example for illustration and for reference. It is a standard definition of a *weak order*. This is a particularly simple structure that occupies a central position in the field of measurement theory and it will provide a basis for much of the discussion in later chapters. Its importance lies in the fact that the existence of a weak order is generally supposed to be a minimum condition for the existence of any measurable property. The definition is quoted (with minor notational changes) from Krantz *et al*, (1971, p.14).

Def.1.2 Let A be a set and $\succsim$ be a binary relation on A, i.e. $\succsim$ is a subset of $A \times A$. The relational structure $\langle A, \succsim \rangle$ is a *weak order* iff for all $x, y, z \in A$, the following axioms are satisfied:

1 Either $x \succsim y$ or $y \succsim x$.
2 If $x \succsim y$ and $y \succsim z$ then $x \succsim z$.

(The name "weak order" is not completely standard but appears to be established in this area. An alternative is "pre-order".)

The axioms express rules concerning the outcome of the measurement procedures and whether or not the axioms are satisfied in a particular context is a matter of empirical investigation. The relational structure and the axioms together are referred to as a *measurement structure*.

Axiom 1 requires that for any two arbitrarily chosen objects in A at least one of the two stands in the relation $\succsim$ to the other. This includes the case where the two are identical, that is the axiom implies that $\succsim$ is reflexive. The condition is usually referred to as *connectedness* although some authors, e.g. Suppes (1957, p.216), refer to it as *strong connectedness*, and reserve the term "connectedness" for the weaker condition:

1' Either $x \succsim y$ or $y \succsim x$ or $x = y$,

which, unlike 1, does not imply reflexivity.

Incidentally connectedness does not in general hold for either of the components $\sim$ and $\succ$ of $\succsim$. Thus the incorporation of $\succsim$ as primitive in the measurement structure brings some economy and technical advantage, albeit at the cost of some complexity in the interpretation.

Axiom 2 is the condition of transitivity. Its interpretation in the context of our example of weight is quite straightforward and it can be seen to be satisfied in an obvious way.

1.1.3 Homomorphisms

The third stage is the construction of a homomorphism from the measurement structure to an appropriate numerical relational structure. This itself falls into two parts, which we shall describe in turn.

The first step is to demonstrate the existence of a homomorphism. The result is expressed in a *representation theorem* of which the following is a typical example.

Th.1.1 Suppose that A is a countable set. If there exists a relation $\succsim$ on A such that $\langle A, \succsim\rangle$ is a weak order, then there exists a function n from A onto Re such that for all x, y $\in$ A,

$$x \succsim y \quad \text{iff} \quad n(x) \geq n(y).$$

The proof of this representation theorem is straightforward and is given in Krantz *et al* (1971, p.15). The theorem is a formal statement of the fact that the empirical structure to which it refers does characterize a measurable property. It expresses the fact that there is a correspondence between the empirical structure and the numerical structure. Empirical relations among the physical objects can be projected on to relations among numbers.

The second step concerns the range of different possible ways in which the assignment of numbers may be made in accordance with the representation theorem. This is expressed mathematically, in a so-called *uniqueness theorem*, as follows:

Th.1.2 If n is a homomorphism from the structure $\langle A, \succsim\rangle$ to the structure $\langle Re, \geq,\rangle$ that has the property described in Th.1.1, and n' is another real-valued function on A, then n' has the same property iff there is a strictly increasing function f from Re to Re such that, for all x $\in$ A

$$n'(x) = f[n(x)].$$

This equation may be viewed as a transformation from one scale to another.

On the first x has the value n(x), on the second n'(x). The importance of a uniqueness theorem is that it governs the *type* of transformations that may occur in the scales associated with a given measurement structure, and it is possible to classify the scales accordingly. Th.1.2 allows the most liberal type of transformation to be found in measurement, and scales to which it applies are referred to as *ordinal* scales.

The type of homomorphism is governed by the type of structure, and different types lead to the different types of measurement scale. For illustration and for future reference I shall complete this sketch of standard formal theory by stating a typical definition of a structure of extensive measurement, along with the corresponding representation and uniqueness theorems.

1.1.4 Example of an Extensive Structure

The following definition is quoted (with minor changes) from Krantz *et al* (1971, p.73).

Def.1.3 Let A be a nonempty set, $\succsim$ a binary relation on A, and o a closed binary operation on A. The triple $\langle A, \succsim, o \rangle$ is a *closed positive extensive structure* iff the following axioms are satisfied for all x, y, z $\in$ A:

1 $\succsim$ is a reflexive, transitive and connected relation.
2 $x o (y o z) \sim (x o y) o z$.
3a $x \succsim y$ iff $x o z \succsim y o z$.
3b $x \succsim y$ iff $z o x \succsim z o y$.
4 $x o y \succ x$.
5 If $x \succ y$, then for any v, w $\in$ A, there exists a positive integer n such that $nx o v \succsim ny o w$, where nx is defined inductively as: $1x = x$, $(n + 1)x = nx o x$. [i.e. nx results from the concatenation of n replicas of x.]

Axiom 1 combines both axioms of Def.1.2 and states the conditions for the structure $\langle A, \succsim \rangle$, derived from $\langle A, \succsim, o \rangle$ in an obvious way, to be a weak order. Axiom 2 is the condition of *weak associativity* of o (relative to $\succsim$). Axioms 3a and 3b are cancellation axioms, Axiom 4 is a positivity axiom. The significance of these axioms will be discussed in some detail in Chapter 7.

Axiom 5 is a so-called Archimedean axiom whose purpose is to ensure that any two objects in A are commensurable. Various alternative forms have been proposed, as for example in Holman (1969). An extended discussion of the role of this axiom together with a treatment of extensive structures in which the Archimedean assumption is dropped is to be found in Narens (1973). The axiom is rather different in character from the others in that it involves existence assumptions and generally places strong constraints on the membership of the set A. Axioms 1 to 4 merely impose conditions on the relation that any object has to any other objects there happen to be. Axiom 5, by contrast, imposes conditions concerning what other objects there *must* be. It is commonly referred to as a *structural axiom* and the apparent need for axioms of this kind in certain structures is of fundamental significance. It is related to the question of the extent to which the definition of a given measurable property depends upon the existence of other objects that share the same property. This question emerges at a number of places in the discussion in the following chapters.

The numerical structure standardly taken to correspond with the empirical structure $\langle A, \succsim, o\rangle$ is $\langle Re, \geq, +\rangle$. The representation theorem for this structure reads:

Th.1.3 Suppose that A is a nonempty set. If there exists a binary relation $\succsim$ on A and a closed binary operation o on A such that $\langle A, \succsim, o\rangle$ is a closed positive extensive structure, then there exists a function n from A onto Re such that for all x, y $\in$ A,

$$x \succsim y \text{ iff } n(x) \geq n(y).$$
$$n(x) > 0.$$
$$n(x \circ y) = n(x) + n(y).$$

The uniqueness theorem reads:

Th.1.4 If n is a homomorphism from the structure $\langle A, \succsim, o\rangle$ to the structure $\langle Re, \geq, +\rangle$ that has the properties described in Th.1.1, and n′ is another real-valued function on A, then n′ has the same properties iff there exists $\alpha > 0$ such that, for all x $\in$ A:

$$n'(x) = \alpha n(x).$$

Scales for which Th.1.4 determines the permissible transformations are referred to as *ratio scales*.

This completes the sketch of the formal theory. I am now in a position to indicate the main problems to be tackled in the thesis.

1.2 Summary of Chapters 2, 3 and 4

In Chapters 2, 3 and 4 we take up the subject of the link between objects and numbers. Each deals with some material relating to the construction of *scales* of measurement.

1.2.1 Chapter 2

In Chapter 2 I shall examine the construction of ratio scales, i.e. scales applicable to the measurement of extensive properties. I shall consider this subject in the light of the question that is to provide the main subject of Chapter 3, namely the question of whether or not extensive measurement is a method of *counting*. I hope to show that it is possible to give an account of scale construction that strongly supports the idea that it is. From this point of view therefore Chapter 2 is a preliminary to Chapter 3, but I think that some of the material is of interest in its own right. Crucial to my argument is a suggestion about the way in which construction of an extensive scale is affected by *experimental error*. The existence of error is a major problem for the theory of measurement. Few comprehensive treatments have been attempted. We shall have occasion to make substantial reference, in Chapter 6, to the theory of *semiorders* due to Luce (1956) which is successful in dealing with certain aspects of error. A different approach in which error is incorporated as a necessary condition for measurement is to be found in Kyburg (1984). While the account of I give of scale construction conforms with the standard picture of extensive measurement, and does not, I think, bring to light anything new at a fundamental level, the treatment of experimental error that is included there may contain some novel material. I end the chapter with an exploration of some analogies between the effects of error and those of intrinsic quantization of physical properties.

1.2.2 Chapter 3

Chapter 3 is an examination of the view that extensive measurement can be construed as a method of counting. The question is obviously of considerable significance from the point of view of our understanding of the application of mathematics to the world. It brings us to the borders of the philosophy of mathematics and, though I emphatically do not wish to stray into that territory, I think that an answer to the question would make an important contribution towards preparing a map for someone who does. We can illustrate the point in this way.

Consider the following two statements.

(i) 2 apples added to 3 apples gives 5 apples.

(ii) A body of mass 2 kg added to a body of mass 3 kg gives a body of mass 5 kg.

These are examples of applications of arithmetic. Anyone who opposes the view that measurement is counting is committed to differing accounts of the two. Ellis, for example, draws a distinction between *primary* and *secondary* applications of arithmetic, of which, for him, (i) and (ii) respectively are examples. Primary applications depend only on the possibility of *enumeration*; that is, he argues, they depend on the very condition which allows us to construct a formal system of arithmetic in the first place. The *existence* of arithmetic is a sufficient condition for primary application. Secondary application on the other hand presupposes stronger conditions, namely those which are necessary for the construction of measurement scales. These happen to be satisfied in the world in which we live, but, Ellis argues, we can imagine a world in which arithmetic and therefore counting are possible but measurement is not (1966,p.20). On this view statement (i) is much closer in logical status to the statement:

(iii) 2 + 3 = 5

than it is to (ii). If this is right it clearly poses a more complex problem for the philosophy of applied mathematics than would exist in the absence of the distinction. Thus there is little doubt that the question is of some importance. What appears in Chapter 3 is not intended as a comprehensive argument in favour of the the view that measurement is counting. Rather it is a detailed examination of some objections to the idea that are directly related to the characteristics of scales. In particular I consider some important arguments due to Ellis based on claims about the conventional nature of scale construction.

1.2.3 Chapter 4

Chapter 4 contains a discussion of other types of scale beside those for extensive properties. The analysis, in Chapter 2, of the construction of ratio scales suggests a different basis for classifying scales from those of existing classification schemes. Briefly, the idea is that scales can be classified in terms of *the kind of information that is incorporated in the construction of the scale*. My intention in Chapter 4 is to examine the possibility of applying this method to the range of types of scale found in physical measurement. It has not been possible to develop the material as thoroughly or as systematically as a full investigation would obviously need. The treatment here is in the form of an exploratory survey of important types of scale.

1.3 Summary of Chapters 5, 6 and 7

The second part of the thesis, Chapters 5, 6 and 7, deals with problems more explicitly related to formal theories. It is clear from the literature of the field that, at least in broad outline, the programme of devising suitable formal structures has been largely successful. The shape of the foundations of measurement has become fairly clear, and something recognizable as a generally accepted standard basic theory has emerged. However it is also clear that even at a fundamental level significant problems of detail remain. In these chapters I shall be concerned with problems that fall into two groups.

One important group is due to the fact that at certain points the correspondence between the properties of physical objects and the properties of numbers breaks down. Physical objects cannot, of their very nature, be manipulated as freely as can numbers and this places significant constraints on empirical relations, constraints which do not occur in the case of numerical relations. Two examples will serve to illustrate this point.

(i) As we noted above, the relation $\succsim$ is usually taken to correspond to the numerical relation $\geq$. Now $\geq$ is always connected in any set of real numbers, but whether or not $\succsim$ is connected in a set of physical objects is by no means so straightforward. For instance, if the interpretation of $x \succsim y$ requires the objects x and y both to be present at the same time in some experimental arrangement, as with the example of weights on a balance mentioned above, there is a problem if x and y refer to a single object. It is impossible to put the same body on both pans of a balance at once, and so whereas the numerical expression $3 \succsim 3$ is perfectly clear an expression such as $x \succsim x$ is far less so. The difficulty becomes more serious when we take into account examples where members of the set A are *states* of physical bodies rather than simply the bodies themselves. Consider temperature measurement where $\succsim$ is usually interpreted in terms of heat flow between two bodies in thermal contact. If x and y refer to different temperature states of the *same* body, so that they are necessarily incapable of being in thermal contact, there is a problem about the expression $x \succsim y$. We shall refer to this as the *problem of noncomparability*.

(ii) A related point arises in connection with the concatenation of physical objects with regard to extensive properties. o is taken to correspond to +. Now addition in arithmetic is closed so that for instance the operation on the number 3 that gives 3 + 3 is quite permissible. The physical concatenation operation, by contrast, cannot be closed, since a body cannot be concatenated with itself, nor with another body that has some constituents in common with it. Expressions such as $x \circ x$, $x \circ (x \circ y)$ and $(x \circ y) \circ (x \circ z)$ cause difficulty. We shall call this the *problem of nonconstructibility*.

These departures from the pattern of numerical relations, both noncomparability and nonconstructibility, lead to difficulties in the formal theory. The two standard structures, the weak order of Def.1.2 and the positive extensive structure of Def.1.3 do not apply immediately in the presence of these phenomena. Significant adjustment is called for in the axioms or in the interpretation of expressions of the language, or in both, and the material of Chapters 5 and 7 is an investigation of matters of this kind. The problems clearly originate in fundamental differences in nature between physical objects and numbers. The consequent failure of the formal structures can be dubbed *essential failure*. However I shall not be concerned with questions of ontology. I do not think that the solutions I shall discuss are bound to any particular theory about the ontological status either of physical objects or of numbers. I shall take for granted the kinds of restrictions on the behaviour physical objects that have been alluded to and examine ways of accommodating them in the theory, and I shall simply assume that numbers are available, whatever their nature might be, to allocate to objects to suit the requirements of a scale. The kind of question I am interested in is this. How does the fact that there *are* differences between the two sets of relations affect our understanding of the nature of the measurable properties involved?

For the other group of difficulties we return to the problem of experimental error. The existence of error gives rise to severe problems in the formulation of formal theories. Generally speaking the axioms of the structure of Def.1.2, the basic conditions for a weak order, fail in the presence of pervasive kinds of error such as limits to the sensitivity of comparison operations. We shall refer to this kind of failure as *experimental* failure. Some theories that incorporate some such aspects of error have been produced, most notably the theory of *semiorders* due to Luce (1956) to which we have already referred above. The treatment that I give in Chapter 6 is developed from the theory of semiorders. One motive for dealing with the problem in this context is that it is possible to indicate interesting parallels between essential and experimental failure. This echos a theme that first appears in the discussion at the end of Chapter 2. The outcome of that discussion is that although intuitively we believe it possible to distinguish between the role of limitations in the measurement procedures and the role of the fundamental characteristics of the measured objects in determining the outcome of the measurement, this distinction cannot be formulated in terms of the measurement operations alone.

I give a brief synopsis of the contents of each of the final three chapters.

1.3.1 Chapter 5

Chapter 5 deals with noncomparability and the failure of the axioms for a weak order. There is a discussion of the consequent problem of interpreting expressions of the formal language in terms of operational procedures, and of adjustments in the formal theory needed to meet the problem. The principal new result is the definition (in Section 5.7) of a structure that I call a *semi-connected order* that offers what appears to be an adequate solution to the problem. The chapter ends with some comment about the significance of the problem of noncomparability and of the proposed solution for our understanding of the comparative terms that are associated with measurable properties.

1.3.2 Chapter 6

In Chapter 6 a simple model for systematic error is suggested that leads again to failure of the weak order axioms. The solution to this problem provided by the theory of semiorders is discussed, and then a possible line of development of this theory is examined. The new results presented in this chapter are the definitions of a *demisemiorder* (Section 6.4) and of a *mutual order* (Section 6.5) both of which are weaker structures than the semiorder and are therefore more generally applicable. The chapter ends with a survey of possible applications.

1.3.3 Chapter 7

The final chapter is given to a discussion of the problems of noncomparability and nonconstructibility in extensive structures. The definitions of two new structures are given. One is a *partly connected order* (Section 7.2.2). This has affinities with the semi-connected order defined in Chapter 5 but is different in scope, and provides a solution to the problem of noncomparability for an extensive structure. The second is a *proper concatenation structure* (Section 7.3.2) which deals with the problem of nonconstructibility. These are incorporated into the definition of a positive extensive structure (Def.7.6) that meets both problems.

In the structure of Def.1.3 above the concatenation operation satisfies the conditions of weak associativity and weak commutativity. These are distinct from the (strong) conditions of associativity and commutativity, in that they merely require pairs of objects such as x o (y o z) and (x o y) o z or such as x o y and y o x to be equivalent, not identical. In the course of formulating the definition of a proper concatenation structure however it is observed that a certain degree of simplification ensues if the strong conditions can be assumed to hold. In particular it is possible to reformulate the structure in Boolean terms. Section 7.4 contains a discussion of the nature of the concatenation operation in a wide range of examples of extensive measurement, with the aim of assessing the significance of the choice between the two sets of conditions. It is concluded that nothing essential to the notion of extensive measurement is lost if the strong conditions are adopted.

This paves the way for the material of the final sections of the chapter. This is to do with a further problem due to nonconstructibility. This is a problem which first came to light in the context of probability measurement (Kraft, Pratt and Seidenberg, 1959) but which is significant for extensive measurement in general. Kraft *et al* demonstrate that cancellation axioms similar to those in our Def.1.3, *unless they are conjoined with strong structural assumptions*, are too weak to generate a certain class of results that we would be inclined to expect the theory to deliver without recourse to any structural assumptions at all. The problem is explained fully in Section 7.6 and a solution, based on a Boolean structure, is proposed in Section 7.7.

CHAPTER TWO

EXTENSIVE MEASUREMENT, QUANTIZATION and EXPERIMENTAL ERROR

2.0 Introduction

In this chapter we discuss the basis of extensive measurement and in particular the construction of ratio scales, which are invariably employed in practice for the measurement of extensive properties. The account is framed in the light of the question that provides the main subject of our next chapter, namely the question of whether or not extensive measurement is a method of counting. I hope to show that the account lends strong support to the view that is.

I begin the analysis of scale construction by describing in Section 2.1 an ideal procedure for what might be termed *ab initio* calibration. This is a procedure for assigning values to a fixed finite set of objects, without recourse to any previously established calibration. We shall take as an example the *ab initio* calibration of a fixed set of weights with the use of a beam balance. The account includes a specification of optimum conditions for calibration, in terms of the membership of the set of objects in question. When these conditions are satisfied the set can be ordered in a particular pattern, which I have called a "well determined order". Examination of the properties of such an order provides strong support for the view of extensive measurement for which we are intend to argue. The concatenation operation plays a crucial role in this, and as a means of highlighting this I preface the account with a brief reminder of how the calibration would be curtailed if, as an artificial measure, we were to forbid concatenations of objects in the operational tests.

In Section 2.1 we treat *ab initio* calibration as a modest self-contained enterprise in that it takes no account of objects not in the set. The construction of a well determined order is dealt with in this section. In Section 2.2 we move on to consider the order as the foundation of a scale of measurement. In this context the aim of the calibration is to provide a basis for assigning values to indefinitely many other objects beyond the original set, and we investigate the adequacy of the procedure described in 2.1 for the construction of scales for various types of extensive property. In particular we consider its application to (a) measurement of quantized properties, and to (b) measurement in the context of experimental error. We complete the account with a brief discussion (Section 2.3) comparing the outcome in situations (a) and (b). The main theme of this discussion is this. The results of any measurement are determined by two groups of factors. One group is to do with the nature of the property being measured and the set of objects that possess it. The other comes from the nature of the measuring operation, including any departure from what we construe as ideal conditions. Intuitively we may be inclined to view these two groups of factors as entirely distinct. However the arguments of Section 2.3 will show, I believe, that the distinction is in the last analysis artificial. I shall argue, for example, that the distinction between the concepts of continuity in (some) physical magnitudes and of perfect sensitivity in measurement operations cannot be drawn in operational terms.

2.1 Calibration

2.1.1 Ordinal Calibration

Suppose that we have a large, though finite, fixed set of objects whose weights (according to any scale) are initially unknown. We denote the set by:

$$A_o = \{a, b, c, \ldots\},$$

and it is understood that a, b, c are all physically distinct. We also have at our disposal a well behaved balance. It is convenient to defer an explanation of what a 'well behaved' balance is to a later point in the discussion in order to avoid a description which begs any questions that are under investigation. In the early stages of the account, it is sufficient to rely on a common understanding of the idea of a normal balance.

Eventually we shall require the balance to be sufficiently large and robust to accommodate in its pans as many of the objects at one time as we care to place in them, but to begin with we shall assume that each pan can take only one at a time. Under this restriction we carry out the following restricted calibration programme. There are two stages.

In stage 1 we compare objects pairwise, placing them in opposite pans, and noting the behaviour of the balance. The whole set is examined in this way, every possible pair being tested and in the light of the outcome the objects are given positions in an order. Intuitively we recognize that the ordering is by weight, but it can be achieved without benefit of any such intuitive insight by application of a simple rule. For any two bodies x and y subjected to this test, if the pan containing x descends then x is assigned to a higher position than y; if the pan containing y descends then y is assigned to a higher position than x; otherwise x and y are assigned to the same position. The outcome can be imagined to be a line of all the objects ranged from the lightest to the heaviest (as we intuitively know them to be) with the complication that at some, or perhaps all, positions there may be two or more bodies piled together. We could represent a typical arrangement this way:

(i)

a,	d,	e,	g,	h,	k,	etc.,
b,		f,		i,	l,	
c,				j,		

This is a *weak* order. It can be regarded as a *simple* order of equivalence classes, and it will be convenient in much of what follows to be able to refer directly to such classes. For that purpose we use the following notation to denote a general sequence of classes in ascending order:

$$C_1, C_2, C_3, C_4, \ldots$$

Using the formal terminology introduced in Chapter 1 we may describe the results of the analysis in this way. The outcomes of the tests define a binary relation $\succsim$ on the set A_0, where intuitively $x \succsim y$ is understood to mean "object x is at least as heavy as object y"; that is, they establish an empirical relational structure $\langle A_o, \succsim \rangle$. The fact that the order in example (i) has been determined unambiguously shows that this structure satisfies the axioms for a weak order which were stated in Def.1.1.

Stage 2 of the calibration procedure is the assignment of numbers, and the essential feature is the rule or rules used for the assignment. One rule is common to all types of calibration and simply ensures that objects are given values according to their positions in the order.

I For any pair of objects, x and y, the corresponding numbers n(x) and n(y) are such that:

(a) if x is at a higher position in the order than y then $n(x) > n(y)$;

(b) if x and y share the position then $n(x) = n(y)$.

To put it informally in terms of our existing notion of weight, the heavier the body the larger the number.

Assignment in accordance with this rule alone produces an ordinal calibration. Any ordered sequence of numbers whatever can be selected to represent the sequence C_1, C_2, C_3,.... Within the constraints of rule I the choice for each equivalence class is arbitrary, and it is a hallmark of this type of calibration that there are as many distinct choices to be made as there are equivalence classes. That is to say in no case is the value for one class uniquely determined by the assignment of values to other classes in the sequence.

A calibration of this sort uses numbers merely as an ordered set of labels. In principle any other ordered set would do, names of Archbishops of Canterbury in historical order, for example, provided that there were enough to match the number of equivalence classes. It is sufficient that a set of labels has the required ordinal properties. The advantage of a numerical representation of course is that infinitely many labels are available whose positions in the order can be recognized systematically. This is important when it comes to treating the calibration as the basis of a *scale* which is to be used to assign weight values to indefinitely many other objects. Even then it is immaterial whether we think of the numerical items attached to the objects as numbers or as numerals. If we opt for numbers then we tag the physical objects with mathematical objects, if numerals then we tag with linguistic objects; both sets of tags have the same ordinal properties.

The outcome of this first calibration programme is not at all satisfactory from the point of view of what we ordinarily understand about the measurement of weight, and in general ordinal scales have little use in physical science, though there are a few examples. One is Moh's scale of hardness in mineralogy, another is Richter's scale for earthquake intensity. We shall say a little more about them in the next chapter. We normally employ scales which allow us to attach more significance to the numerical values than simply that of indicators of position in an order. If two bodies x and y have values $n(x) = 3$ and $n(y) = 6$ on a normal scale of weight, we can deduce more than that y is heavier than x. We are entitled to expect that if x is combined with another body x' to which it is equivalent in weight (perhaps a replica) then the combination will be equivalent to y. y in this obvious sense is equivalent to two of x. Expectations of this sort are at the centre of our intuitive understanding of the concept of an extensive property. At the root of it all is the fact that in constructing an ordinal scale we draw only on information given in the outcome of the operational tests. The construction of any other sort of scale requires the addition of information from elsewhere. In Chapter 3 we discuss this point with respect to a range of types of scale. In the next section of this chapter we pinpoint the source of the extra information required for an extensive property. It is precisely what we have excluded by forbidding concatenation of objects in the tests, and we now turn to the rest of the calibration programme to remedy this omission.

2.1.2 Extensive Calibration

For the complete calibration programme we relax the restriction on concatenation, and assume that we can place as many objects from the set A_o as we wish on a scale pan at the same time. Such a collection is treated as a further composite object. While it is assembled in the pan this object may be tested against some other object, simple or composite, made up from the remaining members of A_o. That is to say the set on which the tests are conducted for this programme is not A_o but a larger set A containing both the simple and the composite objects. A includes A_o as a proper subset. We shall refer to a, b, c,..., the members of A_o, as the atoms of A. Using an obvious notation the composite members are denoted by ab, ac, abc, bc,...etc. and referred to as molecules.

An exhaustive set of tests is again carried out, this time taking account of all possible combinations and pairings, and on the basis of these tests an ordering of A is achieved. In general, deducing the correct order from the the results of the tests will be a more complex task than in the case of the earlier programme. This is because it is not possible to compare all pairs of members of A directly. We cannot compare ab with ac for example. The introduction of concatenation has given rise to a problem that we shall refer to as nonconnectability, and we shall deal with this in formal terms in Chapter 6. However intuitively it is clear that the position of ab relative to ac must be the same as that of b relative to c, The complete order can be deduced in this way, though there is a corresponding minor complication in the fact that it is not possible, as it is in the other simpler programme, to display the completed order by laying out all the objects in line at the same time. For the same reason that it is not possible to compare ab and ac directly it is not possible to display ab at one point in the order and ac simultaneously at another. The order which is the end product of the empirical stage of this programme is a more abstract affair.

As before we can describe the results of the analysis in more formal terms. The concatenation of objects is represented by a binary operation o on A, so that the object formed by combining object y with object x is represented by x o y. Using this terminology we can describe the outcome as the discovery of an empirical relational structure $\langle A, \succsim, o \rangle$ that satisfies conditions sufficient to ensure that $\succsim$ gives a weak order on A.

The second stage of the calibration, as before, is the assignment of numbers to all objects in the order, both atoms and molecules. Again the assignment proceeds in accordance with rule I but the crucial feature of this complete programme is the adoption of a second rule. The function of this second rule is to determine the value to be assigned to a molecule in terms of the values of its constituent atoms. The one that is universally adopted is:

II For any body x that is a composite of separate bodies y and z the corresponding numbers n(x), n(y) and n(z) are such that:

$$n(x) = n(y) + n(z).$$

The point of this rule is perfectly obvious. It ensures that the weight of an object is equal to the sum of the weights of its separate parts. It thus produces an *additive* scale. In a later section I shall discuss the grounds for adopting this particular rule. It corresponds with universal practice in the construction of scales of weight (and of scales for very many other measurable properties, such as length, volume, etc.). The practice reflects, no doubt, a common intuitively held conviction that when two bodies are combined certain of their properties are in some sense added together and that it is natural for the scale to incorporate this basic feature. However some commentators hold that this is an accident of convention, determined largely by the claims of arithmetical convenience. Their view is that there is no essential reason, no reason dictated by the intrinsic character of the property of weight by itself, why an additive scale must be adopted. We shall discuss this question in Section 3.1. However before this I want to consider the effects of applying the rule. The adoption of *any* rule which fixes the values for molecules in terms of the values of constituent atoms has a major consequence for the numerical assignment, and for the purposes of explaining this we shall assume that rule II is accepted.

The major consequence to which I refer is this. It so constrains the assignment of numbers that, *provided that the collection of bodies is sufficiently structured*, once a number has been chosen for one member of the set, the numbers appropriate to all the others are fixed and there is no further choice available. The only arbitrary step is in the choice of the first number. This contrasts very markedly with the situation in programme 1. In order to appreciate exactly how this comes about, and in particular to show what is meant by the phrase "sufficiently structured", it is worth considering some simple examples.

Suppose first that in the original set there are just two distinct bodies a and b. There are two possible situations; they fail to balance (i.e. one, b say, is heavier than the other) or they do balance (i.e. are equal in weight). Consider these in turn.

In the first case the order is:

(ii)

a, b, ab

Notice that ab is in this case assigned to its position in the order not as the result of any test, since no direct test of ab against either a or b is possible, but in anticipation of the application of rule II. Now we may, without loss of generality in the argument, let the first arbitrary choice be n(a) = 1. This places no consequent restriction on n(b) other than that n(b) > 1 as required by rule I. Any of 2, 5, 17, 129 would be acceptable. Here the choice is as open as it would be if calibrating without benefit of concatenation. Of course once n(b) is chosen, then n(ab) is fixed at 1 + n(b) by rule II. But overall there are as many choices as there are atoms in A. We can describe an order for which this is true as *minimally determined*.

In the second case however where the order is:

(iii)

a, ab
b,

there is no choice beyond the first. If n(a) = 1 we must have n(b) = 1, n(ab) = 2. We can descibe an order for which a single arbitrary choice is sufficient, with the help of rule II, to fix all the other values as *fully determined*.

It should be pointed out that in this trivial case the fact that a single choice is sufficient to determine values for the rest of the order is due solely to the fact that a and b are equivalent. This much would have been possible even in the situation considered in programme 1 where addition is not involved. So far rule II has helped only with assigning a value to the additional object ab that concatenation has generated. However with less trivial examples the situation can change.

Suppose now that there are three atoms a,b,c. Various possibilities arise. A minimally determined order is like:

(iv)

a, b, c, ab, ac, bc, abc

In this case the choice of n(a), n(b) and n(c) is as arbitrary as before; again there are as many choices as there are atoms. However there are possible fully determined cases admitting of a single choice only. One is:

(v)

a,	ab,	abc
b,	ac,	
c,	bc,	

A more interesting case is:

(vi)

a,	c,	ac,	abc
b,	ab,	bc,	

In this last example the membership of C_2 is an important link in the calibration. If we put n(a) = n(b) = 1, we then have n(c) = n(ab) = 2 and the values for remaining terms are correspondingly determined. Without some link between a, b and c like c ~ ab, or c ~ a, c ~ b as in (v), the determination could not have been so complete and an intermediate level between minimal and full determination would result as in this case:

(vii)

a,	ab,	c,	ac,	abc
b,			bc,	

Here fixing n(a) also fixes n(b) and n(ab), but n(c) remains open to choice as before. In this case c and the composites which include it float free, as it were, of that part of the system which involves a and b only. C_3 can be related numerically to C_1 and C_2 only as in an ordinal scale.

It is perhaps worth finishing this section by displaying a couple of examples of orders with four atoms a,b,c,d. One possible fully determined order is:

(viii)

a,	c,	d,	abc,	abd,	acd,	abcd
b,	ab,	ac,	ad,	cd,	bcd,	
		bc,	bd,			

and another is:

(ix)

a,	ab,	d,	ad,	abd,	abcd
b,	ac,	abc,	bd,	acd,	
c,	bc,		cd,	bcd,	

It is again readily seen for both of these examples that once n(a) is fixed so is the value for every member of A.

Now despite the fact that the calibration process shown in these examples is highly idealized I think that it indicates some important characteristics of extensive measurement. In the following sections I investigate the properties of fully determined orders and the conditions under which they are produced. I then go on to consider the extent to which they provide an adequate basis for the construction of extensive scales. The idea of setting out these examples has been to expose some of the general features to be found in the structure of a fully determined order and we are now in a position to consider a number of principles that can be abstracted from the examples.

2.1.3 Definition and Properties of a Well Determined Order

The first thing to be established is a condition for an order to be fully determined. With a view to this we first define a set of relations: $\sim_1$, $\sim_2$, $\sim_3$,....etc. The definition is given recursively as follows:

Def.2.1 For all $x, y \in A$ and $n \in N$

$$x \sim_1 y \quad \text{iff} \quad x \sim y$$

$$x \sim_n y \ (n > 1) \quad \text{iff} \quad (\exists u)(\exists v)(\exists m)(u,v \in A \ \& \ m < n \ \& \ u \sim_m y \ \& \ v \sim_{n-m} y \ \& \ x = uv)$$

In the language of a normal scale of weight the expression $x \sim_n y$ means that the ratio of (the weights of) x to y is the integer n. In operational language, however, it may be interpreted very loosely as "x is equivalent to n y's". In some cases this could mean that object x is equivalent to an object formed by concatenating n members of y's equivalence class. However this need not be so. The reason for the complexity of the definition is that we do not wish to assume that there always are n members of y's equivalence class available for this construction. Such an assumption would be unduly restrictive on the membership of A. We do not need to have one thousand individual 1-gramme weights in order to show that a 1-kg weight is equivalent to one thousand of them. To find the ratio of x to y it is sufficient to find a molecule uv equivalent to x for which the ratios of components u and v to y have already been determined, and the definition reflects this.

Inspection of the particular examples of fully determined orders [(iii),(v),(vi),(viii),(ix)] displayed in Section 2.1.2 shows that the following structural condition is satisfied in all cases.

III $(\exists e)(x)(\exists n)(x \sim n\ e)$

This states that there is an object e such that each object in the set is equivalent to an integral number of e's. We can combine Def.2.1 and III to state, less formally, a condition on the membership of the classes C_k.

IV *There is some atom e such that every class C_k includes either e or some molecule of which e is a constituent.*

If any atom answers this description it must belong to the lowest set C_1 of the sequence. Furthermore if one member of C_1 satisfies IV then every member of C_1 does. This is seen in all the examples of fully determined orders of Section 2.1.2. In (ix), for instance, a, b, and c, the members of the lowest class, all occur as a constituents of at least one member of every other class.

It is easy to see that once a value has been chosen for one object in the set C_1 III is sufficient for the assignment of numbers to all other objects (in accordance with I and II) to be unique. This assignment is most obviously achieved by first assuming a value for (the members of) C_1, and then working up through the sequence C_2, C_3, C_4,.... III ensures that each of these higher classes contains a molecule whose constituent atoms have been assigned values lower in the sequence, and then the appropriate value for *all* members of that class is simply the sum of the values of those constituents.

It must be stressed at this point that III is not a necessary condition for an order to be fully determined. To see this consider the following example of an order on a set containing five atoms.

(x)

a,	d,	ab,	ad,	abc,	abd,	ade,	abcd,	abde,	abcde
b,	e,	ac,	ae,	de,	abe,	bde,	abce,	acde,	
c,		bc,	bd,		acd,	cde,		bcde,	
			be,		ace,				
			cd,		bcd,				
			ce,		bce,				

This does not satisfy III but it is fully determined. If, for example, a, b, and c are given the value 1, then, since de ~ abc (see C_5), d and e must be given 1.5. This fixes the values of all the atoms and hence of the entire set. However I do not wish to formulate a more general condition to encompass examples of this kind, since I shall shortly produce reasons for regarding them as incomplete. I shall argue that orders whose structure conforms with III provide the correct basis for an account of extensive measurement. We shall refer to fully determined orders which satisfy III (or IV) as "well determined orders".

One property of a well determined order is that every equivalence class except the highest must have more than one member. (The highest has as its sole member the molecule made up from the entire set of atoms.) In particular this is true of the lowest class C_1 (except in the trivial case where A itself has only one member). Suppose there were only one member of C_1, a say. If this is to be a constituent of some member of the next lowest class C_2, C_2 must contain a molecule ab, say. Then b must be a member of some class lower than C_2. Either, contrary to hypothesis, it belongs to C_1, or else C_2 is not the next lowest class.

An interesting example is provided by a standard set of weights for use with a chemical balance. Typically, such a set will consist of a series of brass weights with values, in grammes:

1, 2, 2, 5 10, 20, 20, 50, 100,.....

These values are chosen to ensure that any value equal to an integral number of grammes (up to the total value of the weights in the series) can be obtained by judicious selection. However such a set cannot be fully determined. Left alone in a laboratory with nothing other than this set of weights (from which we may suppose the value markings have been erased) and a balance, we could not get very far with calibrating them. We could discover the order:

(xi)

(1)	(2a)	(1,2a)	(2a,2b)	(5)	(1,5)	(2a,5)	(1,2a,5).....
	(2b)	(1,2b)		(1,2a,2b)		(2b,5)	(1,2b,5).....

However since there is no way of discovering that (1) is half the weight of (2a) or of (2b), we cannot determine that the order corresponds with values which are multiples of weight (1). The calibration fails to go through because the lowest class contains only one member. The best that can be done to give an additive calibration for this set of objects alone is to remove from the order all except the terms corresponding with multiples of (5) so that we are left with:

(xii)

(5)	(10)	(5,10)	(20a).....
(1,2a,2b)	(1,2a,2b,5)	(1,2a,2b,10)	(20b).....
			(1,2a,2b,5,10).....

This provides an additive calibration for the larger weights in terms of the (5) as unit. The role of (1),(2a) and (2b) is reduced to their being permanently lumped together to provide a second atom equivalent to (5). Alternatively, the situation may be redeemed by using a second 1-gramme weight from another box and the following could then be achieved:

(xiii)

(1)	(1,1')	(1,2a)	(1,1',2a)	(5)	(1,5).....
(1')	(2a)	(1',2a)	(1,1',2b)	(1,2a,2b)	(1',5).....
	(2b)	(1,2b)	(2a,2b)	(1',2a,2b)	(1,1',2a,2b).....
		(1',2b)			

This example indicates a further point of great interest. Suppose that beam balances, instead of working in the familiar way, showed the following behaviour. When objects in opposite pans are equal in weight the beam remains in equilibrium, but when they are unequal it oscillates, with no indication of which object is the heavier. It would still be possible to establish a well determined order such as that displayed in (xiii) by systematic trial and error. The process would be longer and more tedious than before but it could still be done. The outcome of the tests would serve to determine the equivalence classes. The *order of the classes* could then be established by inspection of their members and the application of a simple rule: for any objects x and xy if x is a member of class C_j, and xy is a member of C_k, then C_k is higher in the order than C_j. This shows that an empirical ordering operation is not a necessary requirement for extensive measurement. *Provided that a set of objects does satisfy the conditions for a well determined order an operational test that indicates no more than whether or not two objects are equivalent will suffice.* Thus the role of the ordering operation is not as fundamental to the definition of extensive quantities as has generally been supposed. That balances do in fact indicate by their behaviour which of two objects is the heavier can be seen as a useful empirical discovery about balances. The discovery is exploited to provide a operational *test* of whether the relation "is heavier than" holds between two given objects, but there is no reason to see it as an operational *definition* of the relation. This point has not been widely discussed. The earliest acknowledgement of it that I have seen is in Carnap (1966), where there is a brief informal suggestion that the basis for extensive measurement can be expressed in two rules. One of these is to the effect that there must be an operational equivalence relation $\sim$ such that, for any two objects x and xy where the former is part of the latter, the following holds:

$$\neg(x \sim y).$$

The other rule is our rule II (p. 33 above). A formal treatment is to be found in Holman (1973).

Another significant feature of a well determined order is this. According to III in any class C_i $(i > 1)$ there will be a molecule of construction ef where e is a member of C_1 and f is some other object, atom or molecule which itself must be a member of C_{i-1}. A consequence of this is that a well determined order is regularly spaced. To put it roughly, it goes up in steps equal to e. This in turn fixes the spacing in the sequence of numerical values that results from application of the additive rule II. It is clear that the values assigned to classes C_2, C_3, C_4, etc. must be successive multiples of the value assigned to C_1. (This is not true for the fully determined order of example (x), which, as we pointed out, does not conform with III.) More specifically, if the value assigned to C_1 is 1, the sequence of values for the higher classes must be 2, 3, 4, etc. In this respect the sequence C_1, C_2, C_3,... closely corresponds with the standard series described by Campbell (1920, p.280) and with the standard sequence in Krantz *et al* (1971, p.84), although in those schemes the members of the sequences are objects rather than classes of objects.

2.1.4 The Basis Class and *De Facto* Units

It is apparent that the lowest class C_1 has a special role in the scheme. It provides unit building blocks for the rest of the order. As we have seen there will be at least one member of C_2 constructed from two of these blocks, and more generally there will be at least one member of C_{i+1} constructed by adding one of these blocks to a member of C_i. The idea of extensive measurement as counting comes in very naturally at this point. Determining the value for members of any given class may be thought of in terms of counting these unit blocks. It need not be construed as counting them directly. Indeed with well determined orders characterized as they have been so far direct counting will not usually be possible. In general an arbitrarily chosen member of class C_n will not actually be composed of n members of C_1. It may be a molecule with fewer, larger components or indeed it may be an atom. Rather we may think of the counting in the same sense as in counting money. There we treat individual coins and notes as atoms, and the total value of some amount of cash is found by adding the values of its constituent atoms. The total is the number of unit coins to which the cash is equivalent. It is in this way that the objects in C_1 set a natural unit for the calibration; and measuring any other object amounts to counting the number of these unit objects needed to construct another equivalent to it.

This natural unit is in general distinct from a conventional unit associated with a scale standard. In normal usage in measurement the term "unit" rarely, if ever, refers to the smallest member of a calibrated set of objects. Adoption of the metre as a unit of length, for example, does not restrict us to measuring lengths only in multiples of a metre. This is because it is common practice to assign the number 1 to some class other than C_1 and as a result of this fractional values inevitably arise. Very often the choice is to assign it to that class one of whose members, or subset of members, provides a defining standard for the calibration. But this is simply a matter of convenience and is by no means always followed. A trivial departure is seen in the term "*kilo*gramme" for the standard of mass. Or consider the standard of length for the metric system. For over half a century it was a certain platinum-iridium bar kept in Paris which was assigned the value 1 on the metre scale. In 1961 this bar was replaced as the standard by the wavelength of certain radiation from krypton atoms. At the change this wavelength was assigned not the value 1 but the value $6.05780211 \times 10^{-7}$, on the metre scale. Again this was done simply for convenience. The latter value corresponded, as closely as it was possible to ascertain, to the value under the old standard, and adopting it averted the need for changes to other existing values. To emphasize the nature of the distinction we shall refer to the conventional units, those in which the result is expressed, as *de jure* units, and those which correspond with the class of smallest elements, as *de facto* units. We shall refer to the class C_1 as the *basis class*.

2.2 Scale Construction

In Sections 2.1.1 and 2.1.2 we have viewed *ab initio* calibration simply as a procedure for assigning numerical values to a given finite set of objects without regard for any objects not in the set. We now turn to the matter of scale construction. In this larger enterprise calibration (of some finite set) is meant to provide a basis for assigning values to any object which has the measurable attribute involved, whether or not it is a member of the original set. There are therefore not one but two sets of objects to be considered. The first is the calibrated set itself, which we shall refer to as a *C-set*. The other is the set of all objects to which the scale is supposed to apply, all the candidates for measurement on that scale, irrespective of which are actually measured. We shall call this an *S-set*. There is not in general a unique C-set for a given scale. Indeed once a scale has been established on the basis of some C-set, subsequent measurement of other objects may be seen as a means of enlarging the C-set by the incorporation of new members. On the other hand we may think of an S-set as uniquely determined by the measuring operation on which the scale is defined. It is the largest set A for which, under the relevant interpretation, the structure $\langle A, \succsim, \circ \rangle$ satisfies the appropriate axioms.

The obvious point is that in any given case the two types of set may have different structures. Any C-set is necessarily finite. The total of acts of measurement that have ever been performed and the total of things that have ever been measured on any scale can no doubt be very large but they are finite. An S-set on the other hand can be either finite or infinite. A possible example of a finite S-set is that for mass. It may be the case that all matter is constructed from indivisible fundamental particles, and that there is only a finite number of these particles in the Universe. It is in principle possible for the whole S-set to be calibrated but in most if not all cases of genuine interest it is practically impossible. Thus any C-set is a proper subset of the S-set. An example of an infinite S-set is that for length. We normally assume that there is an infinity of space intervals all of which have length and are eligible to be measured on a scale of length. In the case of an infinite S-set it is of course necessarily true that any C-set is a proper subset.

In the light of this let us stipulate the following condition for the successful construction of a scale of measurement for a specified extensive property.

V *It must be possible to select from the S-set a C-set on which a well determined order C_1, C_2,... can be established such that it is then possible to assign any further object subsequently selected from the S-set to one of the classes C_1, C_2, etc.*

On the face of it the idea that this might be applicable across the whole range of extensive measurement does not look very promising. For a good many properties there will be a significant mismatch between the structure of the S-set and that of any C-set selected from it. There are two necessary (though not sufficient) conditions for establishing a well determined order on a set, namely that (a) the set has a smallest member, and (b) the set of equivalence classes is denumerable. These hold for any C-set, following trivially from the fact that a C-set is finite, but whether or not they hold for the S-set is open to investigation for any particular property. (a) will not be true, for example, if objects can be divided without limit, and if the fragments produced at each division are also assumed to be members of the S-set. For then it will apparently always be possible to select a further object smaller than any member of the current C-set. (b) will not hold in the case of properties whose magnitudes vary continuously. When a rod expands on heating, its length changing from L_1 to L_2 say, it is usual to assume that its length varies continuously within the interval defined by L_1 and L_2. That is to say the rod passes through an infinite number of distinct states whose order type is that of the real numbers. (In many cases both (a) and (b) are thought to fail together but they are independent. The order can be countable without there being a smallest element - the order type may be that of the rationals - and conversely noncountability is compatible with the existence of a smallest element.)

I hope to show that, despite appearances, it is possible to go some considerable distance towards accommodating these differences between the two types of set in the construction of a well determined order. I shall use two general arguments in support of this.

One is, granted that they do exist, the troublesome features of the S-set are expected to show up only under ideal conditions. The occurrence of indefinitely small objects can at most affect the calibration only if the measurement operation is perfectly sensitive. But such conditions are impossible to obtain in practice. The operational tests which are used to construct the order act, as it were, as a selector or filter from the S-set to the C-set, and we can show that limitations in sensitivity or other experimental defects have the effect of distorting or of obscuring certain features of the S-set in the process.

The second argument, if successful, supersedes the first. It is that, given that any C-set is finite, there are features of the structure of the S-set that would be inaccessible to measurement *even under ideal experimental conditions*. We cannot discover in a finite number of tests that in relation to some given property there are indefinitely small objects or that it is capable of continuous variation. This, I shall argue, gives rise to doubts about the intelligibiliy of notions like indefinitely small objects and perfect sensitivity of measuring operations.

First I shall discuss briefly, in 2.2.1, a situation that is free of this problem, that of so-called quantized properties. I shall then deal with the problem of indefinitely small objects and continuity in 2.2.2 to 2.2.4.

2.2.1 Structural and Non-Structural Quantization

For quantized properties condition V appears to hold in a very straightforward way. These properties are supposed to be available only in amounts that are integral values of a fundamental quantum or unit. A prime example is electric charge. This is believed to occur only in multiples of a basic quantum of charge equal to that on the proton or, if we ignore the difference in sign, to that on the electron. The existence of charges of opposite polarity is a complication but it does not spoil the illustration in any way; we can confine our attention to positive charge. According to modern atomic theory any positive charge residing on a body is due to the presence of unbalanced protons. If there are n of them, the total charge is (n x e) where e is the value of the proton charge. The applicability of rule V to this is obvious. Given a suitably sensitive measuring operation, so that differences in magnitude as small as e are detectable, conditions for a well determined order are completely satisfied. There is basis class C_1 whose members are protons and other bodies which bear a single unbalanced proton, and in general the members of class C_n are bodies of charge (n x e). This is a paradigm case. The proton charge is the *de facto* unit, although the *de jure* unit most commonly employed is the coulomb, which is equivalent to 6.241457 x 10^{18} protons. And it is clear that for examples like this the cardinality of the S-set, finite or infinite, is immaterial. The account is the same whether there is a finite or an infinite number of protons in the Universe.

In the case of quantized properties the existence of a class of smallest elements is due to characteristics of the property itself, or at least of the sorts of entity which possess the property. Quantization of charge, if true, is not a product of the measurement operation, but is a fact about charged objects. We could perhaps refer to it in this context as *structural quantization* in that it is a consequence of structural properties of the total measurement system.

We next must consider the case of non-quantized properties. Up to this point in our discussion we have tended to assume ideal experimental conditions. In the case of weight for example this means that the balance that is perfectly symmetrical, perfectly poised, perfectly rigid, free from disturbing influences such as draughts or vibration of the support, and so on. Departures from such conditions always occur in practice, and in general they have two sorts of effects. One is the setting of limits to the sensitivity of the weighing operation. There is a lower limit to differences in magnitude to which the balance can respond. The other is instability. For some pairs of objects a comparison test does not always yield the same result. I now wish to argue that in the absence of structural quantization these nonstructural features of the system give rise to a virtual or *nonstructural* quantization. I shall consider each type of defect, lack of sensitivity, and lack of stability, in turn.

2.2.2 Sensitivity, Precision and *De Facto* Units

The existence of limits to sensitivity is a universal feature of comparison procedures found throughout the whole field of measurement. They are first of all, a characteristic of direct perceptual judgments. If, for example, two bodies are too close in weight, we cannot judge reliably which of them is heavier than the other from how things feel as we hold one in each hand. Or again there are limits to how small a difference in length can be detected by eye, due to limits in how fine a lack of coincidence between end points of rods or similar markers can be recognized. There are limits to how small a difference in intensity of two sounds can be detected by ear. Examples can be cited indefinitely. The use of equipment, such as a balance, or a micrometer screw gauge, or a sound level meter, can increase our powers of discrimination quite considerably, and indeed it is the aim of doing so which very often motivates the experimental scientist in designing apparatus, or devising new ways of making measurements. Nevertheless, however precisely made or however sophisticated in design such equipment may be, limitations always exist. (This is so of course even with the case of a quantized property. We assumed perfect sensitivity in discussing that, but this was merely to simplify the account and it is not needed even there. We need assume only that the measurement operation is sufficiently sensitive to discriminate between elements differing by a single quantum. Experimental tests of the quantization of charge became possible with the development of techniques, such as those in Millikan's famous oil drop experiment, which made it feasible to distinguish between charges on oil drops differing by no more than a single electron.)

The effect at the fundamental level of limits to sensitivity is failure of the axioms of the measurement structure. Consider for example the following situation. A set of objects to be calibrated consists of a number of 1 mg weights together with others suitably distributed, so that given a sufficiently sensitive balance, we could obtain a well determined order with the 1 mg class as the basis class. The balance available however is sensitive only to differences of 2 mg or more. In this case the axioms for an extensive structure do not hold for the full set of objects. For example, given objects a of 1 mg, b of 2 mg, and c of 3 mg we will have a ~ b and b ~ c, but not a ~ c, we will have a o b ~ b, and so on. The transitivity and positivity conditions (Axs. 1 and 5 of Def.1.3) fail.

We now imagine an ideal procedure whereby, faced with the situation we have just described, an experimenter tries to recover what he can in the following way. He analyses the results of the operational tests with a view to constructing some other well determined order in place of the one that he had hoped to build with a basis class of 1 mg objects. The whole set of objects is sifted with a view to picking out a sparser set that does satisfy the axioms. In the course of this analysis he discovers that with the balance at his disposal the best that can be done is to construct an order with a basis class of 2 mg objects (some of these made up of pairs of the original 1 mg collection). He discovers that the *de facto* unit appropriate to his experimental arrangement is 2 mg. The key idea is that the experimenter ensures that rule V is satisfied by judicious choice selection from the S-set.

In this way the experimental arrangement imposes a basis class on any C-set, an effect we have referred to as "nonstructural quantization". This counters the threat which comes from the existence of indefinitely small objects. Any object smaller than a member of this imposed basis class, one obtained, say, by splitting one of its members, lies beyond the scope of the calibration procedures. We might have supposed otherwise. It may appear that we could continue assigning values to smaller and smaller objects by splitting a unit object into say two equal fragments and giving each the value 0.5, splitting one of those into two equal fragments of value 0.25, and so on. But this is blocked by the fact there is no way of knowing that the fragments are equal, or that they belong to a lower class at all. The comparison procedure on which the rest of the calibration is based is, by hypothesis, not sensitive enough to allow us to decide this. It may be claimed that we can sometimes know independently that fragments are equivalent; perhaps we can arrange to split symmetrical objects down a plane of symmetry. The reply to this is that any criterion that is used will itself imply some definite comparison procedure. In the example suggested it is presupposed that there is a means of locating the plane of symmetry which will rely on the ability to measure the object in some way. If the comparison procedure involved is different from that on which the original calibration is based then the equivalence relates to a different property. If it is the same (e.g. suppose we are dealing with a scale of length and we claim to be able to locate the centre of a unit rod in order to split it into two equal lengths) then we are in conflict with the supposition that we have already reached the limit of discrimination.

The ideal procedure described above in which the experimenter discovers the *de facto* unit by constructing a well determined order corresponds to the more workaday practices by which he estimates the effect of the limits to sensitivity on the precision of the measurements. We can relate the precision to the existence of a *de facto* unit. Consider the fact that in practice measurements of the same property are often performed to different degrees of precision. Suppose, for example, that the weight, W, of some object is measured by two methods. In the first a standard chemical balance is used and the result is:

(A) $W = 1.246$ g.

In the second a more sophisticated electronic balance is used and the result is now:

(B) $W = 1.24584$ g.

In (A) the result is given, with the usual convention for significant figures, to the nearest milligramme; in (B) to the nearest one-hundredth of a milligramme. In the light of our analysis of a well determined order we can describe the difference between (A) and (B) quite simply. Both of them use the same *de jure* units, but because the measurements they report differ in precision, because the size of the smallest elements that can be discriminated by the measurement procedure is different in the two cases, they differ in their *de facto* units. In (A) the smallest element is 1 mg in size, in (B) 1×10^{-2} mg in size. These smallest elements give the respective *de facto* units.

We can if we choose make the *de facto* units of the measurement explicit by using them as the *de jure* units. In this vein (A) and (B) can be written in the form:

(A') $W = 1246 \quad u_a,$

(B') $W = 124584 \quad u_b,$

where the unit u_a is 1 mg and the unit u_b is 1×10^{-2} mg. That integral values are now obtained merely reflects the fact that using the *de jure* unit as the *de facto* unit of a measurement is equivalent to assigning values 1, 2, 3,... to the classes C_1, C_2, C_3,....

There are occasions where an indication of the limits of precision cannot be given so neatly within the decimal system as they are in (A) and (B). Suppose that with the electronic balance it is possible to read values only to the nearest two hundredths of a milligramme. This could arise for example if it had a scale graduated in tenths of a milligramme since it is common for an observer to be able to interpolate to no better than one fifth of a scale interval. In this case conventions about significant figures are not sufficient to convey what the precision is and it is necessary to indicate it more explicitly. One way is to replace (B) by:

(C) $W = 1.24584 \pm 0.00001$ g.

The value of W is given as an interval, twice as wide as that understood in (B). This latter could have been indicated more explicitly by:

(B'') $W = 1.24583 \pm 0.000005$ g.

In (C) the *de facto* unit is twice that in (B), and denoting it by u_c we can rewrite (C) to give the value of W as an integral multiple of u_c:

(C') $W = 62292 \ u_c.$

2.2.3 Experimental Instability and Error

Instability in experimental arrangements appears to affect the process of measurement rather differently from insensitivity. Typically its effect is that repeated measurements of a putatively fixed quantity yield a spread of readings. Let us return to the example of measuring W. Repeated attempts with the electronic balance could have produced a series of readings like:

(D)	W = 1.24568 g
(E)	W = 1.24592 g
(F)	W = 1.24601 g
(G)	W = 1.24573 g
	etc..........

This variation reflects instability in the conditions of the experiment. It is standard practice to make assumptions about the stochastic nature of this phenomenon, and on the basis of these assumptions to derive a best estimate of the value by statistical analysis. It would be usual for a value of W to be obtained from the mean, and from the standard deviation about the mean, of just such a series of readings as (D), (E).... to give a typical result:

(H) $W = 1.24583 \pm 0.00003$ g

Statement (H) is formally similar to those in (A') and (B') of the last section, but despite this it is usual to regard it as different from them in kind. The bounds in (H) are accorded a different significance from those in the other two in that they are taken to indicate not limits of precision but limits of error. The underlying idea is that in the list

(D), (E).... each reading is in error, in that each deviates from the (unknown) true value of W to a greater or lesser extent. On this view measurement is like shooting at a target, where features of the practical arrangements of the measurement may conspire to spoil our aim, with a consequent scatter of results. It is likely that (H) also fails to state the true value of W but the limits of error indicate bounds within which this true value is believed to lie with a certain probability.

The distinction between precision and error is not always drawn in this way. It is difficult to settle on a standard usage of the terms "precision" and "limits of error" in the literature of error analysis and sometimes they are used interchangeably. (See, for example, the treatment in Barford (1985), Section 1.4.) I am inclined to think that this looseness reflects the fact that the two concepts are much less distinct than the view sketched above might suggest. In particular there is one characteristic that random error has in common with precision that is of central importance for our discussion, namely that it too sets limits to the fineness of discrimination that can be achieved in the measurement operation. We can illustrate his with a simple example.

Suppose that a balance is exposed to draughts. It is reasonable to assume that this will affect the outcome of a test only if the two objects being compared are very close in weight so that the instrument is nearly at the point of balance. We shall assume the following behaviour. If two objects x, y under test differ in weight by more than a fixed amount, w say, then either x tips the balance against y or else y tips it against x but not both; repetitions always yield the same result. If on the other hand they differ by less than w both outcomes are possible; we may expect to obtain them both in the course of repeated tests. This amounts to a failure to discriminate between the two objects. It is possible to remedy the situation to some extent by adopting a criterion in terms of the relative frequencies of the two outcomes. This can reduce the threshold for discrimination, but it cannot eliminate it. Given that only a finite number of tests can be conducted on any pair there will be a residual difference w', which may be much less than w, but nevertheless for which it will not be possible to decide whether one object is heavier than the other. The effect of this is entirely similar to that of failure of discrimination due to insensitivity. The axioms of the formal theory fail in the same way.

The consequences for the existence of a well determined order are the same as well. The experimental defects lead to nonstructural quantization. Once again we may envisage an ideal method of analysis in which an experimenter, faced with the spread of readings such as in the list (D), (E)...., sifts through the collection of objects with a view to finding a set which satisfies the axioms. He discovers that a well determined order can be constructed with a basis set of elements of magnitude w'. This ideal method corresponds at a fundamental level with the statistical analysis that is likely to be undertaken in practice.

We can give another illustration to support the view that lack of sensitivity and lack of stability are alike in their effects. In the following idealized account of a measurement we show how lack of sensitivity can lead to a random spread of results.

Suppose again that the set of objects being calibrated consists of a number of 1 mg weights together with others which, given an ideal balance, would allow us to construct a well determined order with the 1 mg class as the basis class, but that the balance available is sensitive only to differences of 4 mg or more. Suppose further that, in ignorance of the fact that the axioms do not hold on the total set, an experimenter attempts to arrive at a numerical value for an object p which in fact weighs 20 mg by balancing it against some other objects selected from the set, calibrating as he goes along. The value he arrives at will depend very strongly on the particular selection he happens to make. Suppose he chooses a, b, c and d which have the values shown in the second row:

a,	b,	c,	d
1 mg	1 mg	5 mg	10 mg
1	1	2	4

If a and b are counted as unit objects, then the others are awarded values shown in the third row. c is awarded the value 2 since, given the limit of sensitivity of the balance, it balances ab, and d is counted as 4 since it balances abc. But the combination abcd balances p and so p is measured to be 8 units. On the other hand suppose he happens to make a different selection, a, b, e, f, g, h with values as shown in the next array:

a,	b,	e,	f,	g,	h
1 mg	1 mg	2 mg	3 mg	4mg	8mg
1	1	2	4	8	16

The values awarded to these are again shown in the third row. e balances ab, f balances abe, and so on. The combination abefgh balances p which now comes out at 32 units. (We assume that, happily unaware of the dangers, the experimenter carries out only the minimum number of tests so that no discrepancies are revealed). It is simple to show that other selections can lead to yet other values and a wide range of readings is possible. As in real measurement the existence of the source of error may come to light only because repeated readings fail to agree.

All this suggests that random error can be viewed as experimental insensitivity in a different guise. In the light of this we can perhaps give an alternative description of this sort of error. We can attribute variations in readings, such as those in (D),(E),(F), etc., not to malfunctioning of the equipment but to an excess of precision in the numerical language in which we report the measurement over what the measurement process is capable of delivering. We very often fail to match them exactly because of our ignorance of the size of the *de facto* units. In so far as it is appropriate to describe the readings as erroneous at all, the error is due to something more akin to linguistic impropriety than to experimental maladroitness. Very often our ignorance of the appropriate unit is inevitable in advance of obtaining a series of measured values, and we are obliged to carry out the subsequent "error" analysis as a means of distilling information about the *de facto* unit from the measurements themselves. Adding limits as in (H) is a way of blunting the numerical instrument, so correcting the initial mismatch between numerical precision and experimental capability. This puts (H) very much on the same sort of footing as the earlier statements (A') and (B'). No statement of a numerical result is complete if it does not include an indication of the limits of precision, either explicitly or, by a correct use of significant figures, implicitly. Due regard for this is stressed in the training of experimental physicists.

2.3 *Structural* versus *Nonstructural Quantization*

The central point of the discussion of the last few sections is that extensive measurement is made possible by either structural or nonstructural quantization, and it was argued that in every case one or other of these types always occurs. It seems that, of the two, structural quantization fits our account more readily and it is tempting to speculate that all extensive properties are quantized in this way. In addition to electric charge which we took as an example there are other properties such as angular momentum which are currently held to be quantized. Admittedly there are some entities for which the idea of a corpuscular nature is very difficult to accommodate intuitively, perhaps most of all space and time. However since the advent of modern relativistic theories of space and time we are used to the need to be wary of our intuitions in this area. Just as relativity theory has required physicists to face giving up a concept like absolute simultaneity, which from an intuitive point of view may once have seemed inviolable, so quantum theory may require us to face other difficult possibilities. Angular momentum is defined classically in terms of speed of rotation in space, and if it is quantized then it is plausible to suggest that this is a consequence of quantum properties of space and time. For any property whatever there is a possibility that there is an underlying structural quantization beyond the limits of precision of current observational methods.

By the same token, however, with regard to any property that is at present supposed to be quantized there is a corresponding possibility of discovering that what were thought to be quanta can after all be split into smaller amounts. Current theory of the substructure of fundamental particles involves reference to quarks, particles which are supposed to carry charge lower in magnitude than that on the proton, though so far there is no evidence for the existence of free quarks. We clearly cannot stake an argument on the truth of any particular theory of physics. The point is simply that measurement cannot provide decisive evidence either way as to whether or not any given property is (structurally) quantized.

The impossibility of deciding this question is not simply a product of contingent experimental limitations. It lies deeper than that. It could not be settled even if ideal experimental conditions were realizable. The condition that the set A (interpreted here as the S-set) has no smallest member is equivalent to saying that for any object x in A there must be a further object y such that x is greater than y. The following condition must obtain:

VI $$(x)(\exists y)(x \succ y)$$

Since A is infinite VI can neither be confirmed nor disconfirmed in finitely many empirical tests.

For the very same reason, we could not even discover that we were fortunate enough to have a perfectly sensitive experimental arrangement at our disposal. The difficulty about the concept of perfect sensitivity in a measuring operation, as far as we intuitively understand it, is that it cannot be expressed solely in operational terms. VI is a strong condition on the sensitivity of the operation as well as on the membership of A. It is a sufficient condition for perfect sensitivity, but ordinarily it would not be thought of as a necessary one. We might be inclined to suppose that it is logically possible to have a perfectly sensitive instrument even if the set of objects to be measured is insufficiently rich for perfect sensitivity to be required. In this situation the superlative qualities would simply be redundant. However it is important to notice that there is no empirical warrant for the separation of the idea of perfect sensitivity from that of indefinitely small objects. From an operational point of view the two are inextricably linked. I shall elaborate on this point.

We recall the problem that was alluded to at the start of Section 2.1.1, namely that of giving a specification for a well behaved balance. In the interests of getting started on the discussion we set this question aside, and simply relied on an intuitive notion of an ideal balance. At that point we perhaps understand perfect sensitivity, for example, to imply that there is no lower limit to a difference between two weights that can be detected so that two objects will balance, and therefore will be counted as equivalent, only if they are exactly equal in weight. But this goes beyond operational concepts. Use of the expression "exactly equal", clearly intended to convey something other than mere operational equivalence, already presupposes the kind of mathematical representation whose validity is at present under scrutiny. It is intelligible only by virtue of our normal habit of thinking of weight as a quasi-numerical variable. However from observation of the behaviour of the balance we can at most establish that an equivalence relation exists on a certain set of objects. If the appropriate axioms are obeyed we are in a position to describe the balance as well behaved, but well behaved relative to the set of objects in question. We might say that it is *sufficiently sensitive relative to A*. This is the best that can be done from an operational standpoint. So with the term "perfect sensitivity". It can indicate no more than that a measuring operation is sufficiently sensitive relative to a set which has indefinitely small elements. There seems no alternative but to regard VI as a specification of perfect sensitivity just as much as it is of indefinite smallness.

The converse of this is that if VI does not hold for a particular property there is nothing to dictate which of the two components of the system - the set of objects or the measurement operation - should be held responsible. Consider again the example of quantization of charge. The existence of a quantum of charge is usually thought of in terms of the structure (or perhaps lack of structure) of protons or electrons. But all methods of measuring charge are based on the interaction of the measured charge with the electric field of other charges, and the quantization could well be thought of as a function of this interaction. If this line of thought is correct then I think that the distinction that we have freely assumed in the discussion so far, namely that between structural and nonstructural quantization, ultimately may not be sustainable. We must view quantization simply as a feature of the total system.

CHAPTER THREE

EXTENSIVE MEASUREMENT and COUNTING

3.0 Introduction

If the idea that was explored in the last chapter - that well determined orders provide the proper basis for extensive meaurement - is acceptable it clearly lends strong support to the view that extensive measurement is, fundamentally, a process of counting; more precisely, that it is a means of counting unit elements. In this chapter I wish to explore this idea further. In particular there are some important counter arguments that must be considered.

As was indicated in the summary in Chapter I what follows in this chapter is not intended to be a comprehensive argument in favour of the view that measurement is counting, let alone a compelling one. What I wish to do is discuss some objections to the idea that are directly related to the characteristics of measurement scales. There are two in particular that I shall discuss at some length. The first is that the adoption of an additive numerical representation, on which this view of extensive measurement crucially depends, is entirely conventional. The argument is that additivity in the representations, though intuitively appealing, is not dictated by the logical characteristics of measurement and that the fact that we do employ additive scales has come about, as it were, by historical accident. Thus the notion of counting cannot be an essential ingredient of extensive measurement. The second objection is to the effect that although additivity in the representation is a necessary condition for the counting of units, it is not sufficient. The argument is that even if an extensive scale is based on an additive representation, this does not guarantee that an object of magnitude n units on that scale can be described as being composed of n elements of unit magnitude. We shall examine an interesting model of an extensive structure which Ellis describes and proposes as a counterexample.

3.1 Is the Additive Rule Conventional?

3.1.1 Additive and Nonadditive Extensive Scales

In practice all extensive scales are based on additive numerical representations. The formal basis for this was outlined in Section 1.1.4. Provided the appropriate axioms for an extensive structure are satisfied, a homomorphism exists from the empirical structure $\langle A, \succsim, o\rangle$ into the numerical structure $\langle Re, \geq, +\rangle$. The additive rule comes from correlating the physical operation of concatenation with the arithmetical operation of addition. However other kinds of representation are possible. These are based on different numerical structures, in which concatenation is correlated with a different arithmetical operation. A simple example comes from replacing $\langle Re, \geq, +\rangle$ by $\langle Re, \geq, x\rangle$ where x is the usual numerical operation of multiplication. If there is a homomorphism from $\langle A, \succsim, o\rangle$ into $\langle Re, \geq, +\rangle$, then there also exists a homomorphism into the structure $\langle Re, \geq, x\rangle$. We can obtain a numerical assignment based on this if we replace the additive rule II of section 2.1.2 by a corresponding multiplicative rule:

II' For any body x that is a composite of separate bodies y and z the corresponding numbers n(x), n(y) and n(z) are such that:

$$n(x) = n(y) \times n(z).$$

An extensive scale based upon this rule would strike us as rather strange. Consider such a scale for weight. All objects would be assigned a number greater than 1, the number 1 itself now being reserved for the empty pan. If two equal weights were concatenated the combined weight would be the square of the value of each. A 3 kg weight, for instance, (assuming we adopt the same name for a unit), combined with another 3 kg weight would give a composite body of weight 9 kg. A person's weight would increase at a prodigious rate through childhood, and the effects of overeating would appear very alarming to those nurtured on an additive scale. The scale markings on the common spring balance would be spaced logarithmically instead of linearly. We could go on indefinitely sketching in this unfamiliar picture.

Similar things could be said if we adopted this rule for scales of length, of time interval (duration), or of any other extensive quantity. With a multiplicative scale for length, for example, if the distance from A to B is 4 miles, a journey from A to B and return to A is 16 miles in all. Even stranger, on return to A, we are 1 mile from where we started. If a rod 2.25 units in length is cut in the middle, the resulting pieces are each 1.5 units in length. With a multiplicative scale for time interval, an individual will be 1 year old at birth. After a cyle of four seasons (if we choose to define the unit in this way) he will be 2 years of age, after 2 cycles he will be 4, after 3 cycles he will be 8, and so on. It will be correct to describe him as having lived one year longer than his twin brother though he remains the same age as him. If his sister is born when he is 4 years old, it will always be correct to say that he has lived four years longer than her though the gap in their ages starts off at 3 and then widens considerably as time goes by.

Whatever may be the case with details of this sort the central point of the argument we are considering is that as far as the capacity to express and convey quantitative information is concerned a multiplicative scale is just as adequate as an additive scale, and the same applies to a host of other possible scales based on alternative representations. Strange as their characteristics may seem these other scales could be employed consistently. Admittedly, it would be necessary to reformulate numerical laws in physics and other branches of science. Most would be more complicated, some hideously so, though some others might well be simplified. To give just one example of the latter, there are a large number of experimental laws that state that some quantity varies exponentially with time. An important example is the law of radioactive decay. Suppose that a radioactive sample initially contains n_o active atoms. If time t is measured on a conventional additive scale, then the number of active atoms remaining after a time t has elapsed is given by the formula:

$$n = n_o \times e^{-at}$$

where a is a constant dependent on the particular species of radioactive atom. However if time were measured on a multiplicative scale this could be replaced by:

$$n = n_o \times t'^{b},$$

where t′ indicates the time value on the new scale, and b is another constant. This would generally be regarded as mathematically simpler than the first.

On the basis of considerations of this kind some writers regard the adoption of one type of scale rather than another as entirely a matter of convention. There is nothing to compel a particular choice. We may as well decide on grounds of convenience and select those representations that lead to the simplest arithmetic or to the greatest economy in expressing the laws of physics. But these are alleged to be the only possible grounds for a decision. As Ellis puts it "only reasons of mathematical simplicity can guide us in the choice of a fundamental scale" (1966, p.83). There is held to be nothing intrinsically additive about an extensive property like length, or weight. Our attachment to additivity, which makes contemplation of the nonadditive representations described above so disconcerting, is simply a product of lifelong familiarity with traditional methods, and is no reason for awarding additive scales a unique status. To quote Krantz *et al*, "despite its great appeal and universal acceptance, the additive representation is just one of ... infinitely many, equally adequate representations" (1971, p.102).

Now if we are prepared to accept the claim that nonadditive scales are equally adequate, it appears that the idea of extensive measurement as counting can hardly be sustained. For the claim implies that whether or not a 2-kg weight is composed of two 1-kg parts is fixed by convention, and not by anything characteristic of the property of weight. The idea that when we perform a measurement on some extensive property we are counting something is seen as no more than an incidental byproduct of our conventions; we have in effect been misled by the language of measurement that we have chosen to employ. To suppose that a 2-kg weight is referred to in that way because it is composed of two of something is simply a mistake, equivalent perhaps to supposing that the name of a Boeing 747 is intended to indicate the number of components in the aircraft. If in the wake of accepting the claim we were persuaded to adopt nonadditive representations, we would find it necessary to adopt new attitudes towards the numbers involved, and in the process the illusion of counting would disappear.

I think the claim that we cannot count with a nonadditive scale needs more subtle consideration than has been given to it in the last paragraph and I shall remedy this very shortly. But the point I wish to make first is this. Whether the claim is accepted or not the argument from the equivalence of additive and nonadditive representations to the denial that extensive measurement is counting does not work. One way or another it begs the question at issue. If an additive scale is not essential for counting and a nonadditive one will do then there is no argument. If on the other hand an additive scale is essential then the claim that the two types are equally adequate can be made only after it has been established that there is no counting to be done. The argument is circular because in effect it goes like this. If we ignore the need to count anything then nonadditive scales are on an equal footing with additive scales. Therefore there is nothing to count.

What the argument has ignored is the possibility that we begin by recognizing that there is something to count and that we choose additive scales to count with. If we first grant that determining the weight of a body really does amount to counting the number of unit elements into which the body may be decomposed then the assumption of additivity is no more conventional here than it is in any other context in which the arithmetic of counting is used. The assertion that a 2-kg weight added to a 3-kg weight gives a 5-kg weight is conventional to precisely the same extent that the assertion that 2 apples added to 3 apples gives 5 apples is conventional. That is to say I am prepared to admit that there may be conventional elements involved in the choice of a representation but that if so they are not distinctive features of extensive measurement, but are common to counting generally. By way of elaborating on this I wish to take up the point from the last paragraph about whether or not counting presupposes an additive scale.

3.1.2 Additive and Nonadditive Counting Systems

A conventional element in counting is the names of the numbers. We can use Arabic numerals or Roman numerals, we can count in different languages. But a change in the names of the numbers is not a change in the nature of the counting process. We are doing the same thing whether we count in English or in French. This is true even in those cases where the change is simply a reallocation of existing names in the same language. The change from a decimal to a binary system is just such a change, and we can count equally well in both systems. We can imagine more complicated changes. Suppose for example it were decided to replace the sequence of numerals we ordinarily use - those in row (a) - by the alternative set shown in (b).

(a)	0	1	2	3	4	5	....
(b)	1	2	4	8	16	32	....

This system would be more difficult to use. The rule for generating the next numeral after any given one in (b) is certainly more complex than for (a), involving as it does a process which bears a close resemblance to what we know as multiplication by 2. To help them with this, and with addition generally, children brought up under the new regime would have to learn what would no doubt come to be called addition tables. In general rather more mental agility would be called for than we are accustomed to in arithmetic. Nevertheless this new system could be used consistently, and in particular we could count with it perfectly well.

Now the change from (a) to (b) is a change in the system of counting that is entirely analogous to the change from the additive to the multiplicative representation for a scale of weight (i.e. the change from rule II of Section 2.1.2 to rule II′ of this section) that we discussed as an example at the beginning of the section. Alternatively however, it can be described as a change in the language. For anyone reared only in the knowledge of (b), the numerals in that sequence will have the same meaning as the corresponding numerals in (a) have for us. He will be as happy in the belief that he has 32 fingers on each of his 4 hands, as we are in accepting that we have a total of merely 10. From his point of view (b) is in a new language. On the other hand an individual who starts off with (a) as his first numerical language may be more inclined to see (b) in terms of an arithmetical transformation within the same language. At least to begin with he will make regular use of the transformation relation:

$$n(b) = 2^{n(a)}$$

However we can expect that with practice he will rely on this less and less. There will come a time for example when, on being told that the number of guests invited to dinner has gone up from 64 to 128, instead of nervously doing some mental arithmetic, he immediately enquires who the extra person is. As the result of continued use (b) is seen less and less as a transformation of (a) and more as a new numerical language. At the point at which he can dispense with the transformation relation, and mental arithmetic is unnecessary, the process is complete. He has become bilingual and translation is no longer needed. When he reaches the point of using the sequence (b) to count directly, he has in effect pressed the new numerals into the same service as the old and by doing so has converted the sequence into an additive sequence. It has become suitable for counting by virtue of being used for counting. Common, if trivial, examples of this sort of thing occur with methods of scoring in some games. Consider the sequence 15, 30, 40 from tennis. If someone familiar with the game is informed that the score stands at 40-15 he recognizes without effort that 4 points have been played.

Suppose we now label (a) and (b), as well we might initially be inclined to do, as *additive* and *nonadditive* counting systems, in parallel with our examples of additive and nonadditive measurement scales. What is the import of these labels? What distinction do they actually signify? We have seen that both systems can be used for addition, in that they both incorporate suitable operations for the purpose. To put the question another way, how would someone unfamiliar with either system know which way round to put the labels? He might, after some examination, decide that one was simpler to use than the other, though even then his decision might well depend on the characteristics of his own counting system. In any case it is not clear how simplicity of manipulation can be relevant. Would we withhold the epithet "additive" from the Roman system of numerals on the grounds that it is worse than both of them in this respect? I think that the only ground for labelling (b) as nonadditive is that its addition operation is not the same as (a)'s. This is a formal distinction made in a fixed language and renders the notion of additivity entirely relative. We are entitled to state that if the terms of sequences (a) and (b) belong to the same language then sequence (b) is nonadditive relative to (a). Equally we are entitled to say that (a) is nonadditive relative to (b).

The distinction in question is clear only if the terms used in each system have a fixed meaning. It has no meaning otherwise. Thus, at an informal level the distinction is vitiated at the very point at which we start to treat the sequences as counting systems. In the act of using the sequences to count with we promote shifts in meaning which destroy the assumption on which the distinction was based.

I suggest that the position is the same with respect to nonadditive scales for extensive measurement. Formally we can recognize the difference between an additive and a nonadditive scale in terms of the relation to our present numerical system. If it were insisted that for length, say, we adopted the multiplicative scale that was described earlier in this section, we could make consistent use of it. But this does not prejudice the claim that length measurement is counting. After sufficient practice with the new scale we would be inclined to attach to the numbers on the new scale the same significance that went with the numbers they had replaced on the old scale. We would become accustomed to the fact that a rod 25 units long can be split into two pieces (assuming now that "two" has its traditional meaning) 5 units long. We would come to recognize that in this context 25 units is twice as much as 5 units. What started off looking like a new scale, is later revealed as the old one in disguise. In the act of using the new scale to make measurements, we bring about those shifts in meaning which tend to eliminate the distinction we started with. I conclude that the existence of nonadditive scales is no way threat to the view that extensive measurement involves counting.

3.2 Is the Method of Concatenation Conventional?

3.2.1 Inches and Dinches

I now wish to consider the second of the possible objections to the view of extensive measurement as counting. We shall discuss an intriguing example of an extensive structure, first described by Ellis (1960), for which, apparently, the counting idea does not work. It is a structure for length measurement, involving an unorthodox method of concatenating rods. We can best explain the argument with reference to the formal properties of the structure $\langle A, \succsim, o \rangle$. Ellis points out a possible interpretation of the concatenation operation for this structure which is different from the one that is traditionally assumed, which nevertheless satisfies the same axioms, and therefore, he claims, yields an alternative additive measure of length. He then points out that this new scale fails to conform with the picture that we have tried to paint above of extensive measurement as a method of counting unit elements. We first give the details of this example.

His rule of concatenation is this. The object x o y is formed by placing x and y end to end so that the line joining the extremities of x is perpendicular to the line joining the extremities of y. If x and y are both single rods, the object x o y is as shown in Fig.3.1. (Figs.3.1 to 3.7 are on the following page.) If a further rod z is then concatenated with x o y according to this rule so forming the object (x o y) o z the result is as shown in Fig.3.2. It is formed by setting z at right angles to the line joining the ends of x o y. These composite objects may be

Fig.3.1

Fig.3.2

Fig.3.3

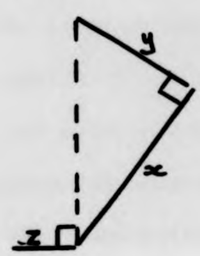


Fig.3.4

Fig.3.5

Fig.3.6

Fig.3.7

compared with others. The interpretation of $\succsim$ in this scheme involves comparing what we might call the major chords of the objects, namely the straight lines between their extremities. This method of comparison is in effect a generalized version of that employed for straight rods alone. For a single rod the major chord lies along the rod itself, for the object in Fig.3.1 the major chord is the hypotenuse LM, for that in Fig.3.2 it is the line LN, and so on.

Ellis's argument depends upon the fact that the new interpretations of $\succsim$ and o endow the structure $\langle A, \succsim\ o \rangle$ with the same formal properties as before. That is to say the axioms of Def.1.3 are still satisfied. It may be thought on initial inspection that difficulty will arise with meeting the requirements of weak associativity and weak commutativity of o. The concatenation operation is sequential in nature and if associativity and commutativity are both thought of in terms of the spatial configurations that the concatenation operation produces it is seen that significant differences arise from differences in ordering. For example, suppose we change the order that leads to the object in Fig.3.2 by first joining y and z to form $y \circ z$, and then joining this object with x to give $x \circ (y \circ z)$. The result is as shown in Fig.3.3. The objects in Figs.3.2 and 3.3 are plainly different in shape. However elementary application of Pythagoras's Theorem shows that their major chords are equal in length. Hence Axiom 2 of Def.1.3 is satisfied. The same is true for commutativity. Suppose that we change the order by first laying down z and then joining the (already formed) $x \circ y$ to it to give $z \circ (x \circ y)$. The result is shown in Fig.3.4. There are in fact six distinct configurations that can be obtained with three rods. The other three are shown in Figs.3.5 to 3.7. And incidentally the variety is richer than these diagrams suggest. They portray planar configurations, but there is no intention to restrict the concatenation operation in this way. For example, in Fig.3.2 rod z may be rotated freely about the axis LM. Thus a term such as (x o y) o z stands for a whole set of three dimensional objects of which the object in Fig.3.2 is a degenerate two dimensional case. Again using Pythagoras's Theorem we may show that the major chords of all objects in Figs. 3.2 to 3.7 are the same length. This is sufficient to meet the requirements of Def.1.3.

Ellis now supposes that a numerical representation is devised for this scheme in an entirely similar way to that for the more orthodox one. One object (or rather an equivalence class of objects) is chosen as the standard, and given a unit name - he suggests 1 dinch (short for "diagonal inch") as an example. An object formed when two dinch-long objects are joined at right angles then has a length of 2 dinches (in the new sense of "length" based as we have stipulated on a comparison of major chords). If a third dinch-long object is joined to this in the prescribed fashion the resulting object is 3 dinches in length, and so on. As well as composite objects, (i.e. molecules), we can of course have straight rods, (i.e. more atoms) of lengths 2 dinches, 3 dinches, and so on. A 2 dinch rod for example is one that just matches the major chord of the object formed by the two 1 dinch rods (which is the same as the diagonal of a 1 dinch square). And similarly for higher values. The relation between values on this scale and those on the normal scale is simple to state. Assuming for simplicity that the same standard is used for each, i.e. that:

$$1 \text{ dinch} = 1 \text{ inch},$$

then we have the relation:

$$n \text{ dinches} = \sqrt{n} \text{ inches}.$$

Thus this method allows the construction of a scale which appears to be an alternative scale of length and which is not linearly related to the usual scales.

We could of course have constructed this new scale on the basis of our normal interpretations of $\succsim$ and o, by producing a fully determined order as before, but then opting for a nonadditive representation. We could simply have assigned the numbers $\sqrt{1}$, $\sqrt{2}$, $\sqrt{3}$, etc. to classes C_1, C_2, C_3, etc. of our original order. [An important detail is that we cannot accommodate all the rods from Ellis's order in this one. If his C_1 is the same as ours then no objects from his classes C_2 or C_3 appear anywhere in our order. Our C_2 is his C_4, our C_3 is his C_9 and so on. There are issues of incommensurability involved here. However I do not believe that they impinge upon the argument at this point.] But the essential point about Ellis's method of construction is that not only are the empirical foundations formally similar to those for the usual scale, but also the numerical assignment is made according to an additive rule. Formally, therefore, the two scales are on an equal footing.

The existence of this new scale now challenges the idea that measuring the length of a rod involves counting unit elements. For the possibility of viewing it this way seems to depend on which scale is adopted. As Ellis points out a straight rod n inches long can be divided into n sections each 1 inch long, but a straight rod n dinches long cannot similarly be divided into n sections each 1 dinch long. For example if rod AB, with midpoint M, is 4 dinches long it is composed of only two 1 dinch sections, AM and MB. A rod 2 dinches long decomposes into two half-dinch rods.

Ellis argues that we cannot escape this consequence simply by rejecting the new scale on grounds of its lack of intuitive appeal. The alternative interpretations of $\succeq$ and o may well seem unnatural and we are certainly more comfortable with the idea that the combined length of two rods is exhibited when they are abutted along a straight line than with that of the rectangular method. But this he dismisses as a "feeling", the result of our thinking having been "coloured by our upbringing". On his view our intuitions about this matter owe much to the psychological conditioning of childhood learning, and nothing to the logical features of measurement, and in no way justify rejection of the new scale as less appropriate to length measurement than the traditional one.

I have two arguments against these conclusions. The first is that although Ellis's new scale may be a legitimate scale there is no reason to suppose that it is alternative to the normal scale of length. The second is that it is not the case that measurements on his scale cannot be construed in terms of counting elements. I shall elaborate on each.

3.2.2 Are dinches units of length?

The first point concerns the question of what property the new scale actually measures. Ellis assumes that it is an alternative scale for length, but it is by no means clear that this is right. What his arguments show is that there are two different interpretations under which the structure $\langle A, \succsim, \circ\rangle$ satisfies the same set of axioms. This is clearly not sufficient for the corresponding measurable properties to be identified with each other since the same thing apples to a whole range of different interpretations which are associated with different extensive properties. If A is a set of rods and $\succsim$ and $\circ$ are interpreted in the standard way in terms of operations with a balance, then $\langle A, \succsim, \circ\rangle$ satisfies the same axioms but we do not deduce that a scale of weight is an alternative to a scale of length. There must be a stronger relation between two interpretations for them to be counted as determining the same property. Would it be sufficient for the two to produce the same ordering of A? This would rule out counting weight and length as the same, since if A contains rods of various thicknesses or made of various materials, the weight order and length order would in general be different. However it would not exclude other pairs of properties that we recognize as different. If A is a set of pendulums and if $\succsim$ and $\circ$ are interpreted in the usual way for (i) length measurement, and (ii) time measurement, then the ordering of A is the same for both cases. And yet length is not time.

It is perhaps in the spirit of Ellis's discussion, though I do not think that it is made explicit, that the two interpretations in question differ only in the significance attached to the concatenation term o and that $\succsim$ is the same for both cases. The idea could be that it is this term, or the interpretation of it, it is the ordering relation alone, that fixes the property. The nature of the concatenation operation is immaterial provided that it leaves the axioms satisfied. However, even supposing this point is granted, the claim that $\succsim$ has the same meaning in the examples we are comparing is dubious. Let us denote $\succsim$ in the standard system, based on linear concatenation, by $\succsim_s$, and that in the Ellis's system by $\succsim_E$. As we saw $\succsim_E$ involves comparing major chords, the straight line distance between extremities. Now when both of the objects under comparison are single straight rods the methods of $\succsim_E$ and $\succsim_s$ are equivalent but this does not give them the same meaning. For suppose that we keep to the rectangular method of concatenation but adopt as an alternative to $\succsim_E$ the method of comparing objects which is illustrated in the following example. Two objects such as x and y in Fig.3.8 (Figs.3.8 to 3.10 on next page) below are set with extreme points L, L' in coincidence and sections LM, L'M' along the same axis. x is then rotated about M until the position in Fig.3.9 is reached. Then y is rotated about M' to give the position of Fig.3.10. In this final position N extends beyond N' and this is expressed by $x \succsim_F y$. It would be tedious to formulate a general rule, but the sense of $\succsim_F$ is clear from this example. It amounts to a comparison of the combined lengths - lengths in the orthodox sense - of the component rods. The main point is that $\succsim_F$ as well as $\succsim_E$ is equivalent to $\succsim_s$ when two straight rods

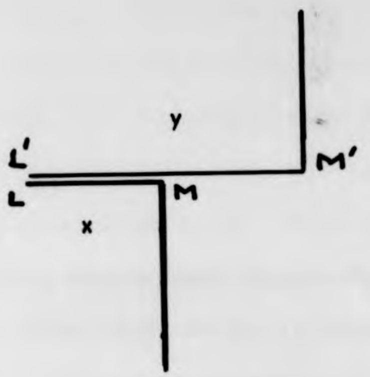


Fig.3.8

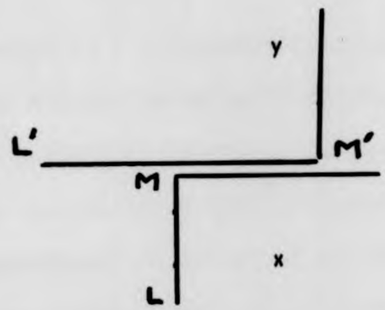


Fig.3.9

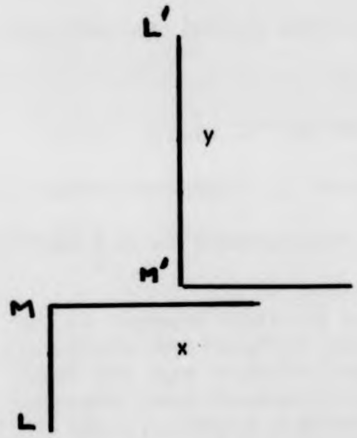


Fig.3.10

are compared. But $\succsim_F$ is certainly not the same relation as $\succsim_E$. I conclude that there are no grounds for regarding $\succsim_E$ as the same as $\succsim_o$ and therefore no grounds for supposing that they order A according to the same property. It may well be that straight rods have some property, hitherto unnoticed but now revealed by $\succsim_E$, which increases monotonically with their length, but there is no more reason to identify this new property with their length than there is to identify the area of a circle with its diameter.

3.2.3 <u>Can dinches be counted?</u>

The points made so far in this section do not completely dispatch the difficulties which appear to stem from Ellis's example. There still is the problem that whether or not the dinch scale is a scale of length it is a scale for some extensive property. It still may be a counter-example therefore to the claim that extensive measurement is counting. We still face the problem that although the dinch scale has impeccably additive foundations, we cannot split a 2 dinch rod into two 1 dinch rods.

Ellis argues on the basis of this example that the view that extensive measurement is counting is mistaken. He attributes the view to a misconception of what a unit is.

> It is thought that if one object possesses a unit of a certain quantity and another object also possesses a unit of that quantity, then the two objects together must possess two units......Thus it is thought that fundamental measurement is simply a matter of counting units......[But] a physical object, or group of physical objects cannot be said to contain so many units of length (inches, say) in the sense that a basket can be said to contain so many eggs.
>
> (1966,p.85)

My reply to this is that we do not claim to be counting units. We are counting objects, unit objects. When we measure the length of a rod we are not counting millimetres, we are counting millimetre-long bits of rod. When we weigh a lump of copper we are not counting milligrammes, we are counting milligramme-sized pieces of copper. The technique is well understood in banks, where it is common practice to use a balance to count copper coins in a bag.

The fact still remains that if we measure the length of a straight rod to be 2 dinches, there are not two 1 dinch sections in the rod to be counted up. One way of dealing with this would be to concede that there are not two unit elements in the rod itself, and to accept that the measurement does not count bits of the rod. Instead we could be content with the claim that the measurement counts the number of unit rods that would be needed to construct an object equal in length to the one being measured. That is straightforward. A 2 dinch rod is equal in length to two 1 dinch rods joined at right angles. Ellis suggests that this is not satisfactory since the L-shaped structure formed by the combination is not itself a straight rod. It is not an object "of the same sort". I do not think that he is entitled to this complaint, however. According to his account both the straight rod and the L-shaped structure are members of the set A, in which indefinitely many other shapes are also represented, and they both have length. This is sameness of sort enough. To find the diversity of shape objectionable is surely to question the legitimacy of the concatenation operation that produces it. When devising a scale of weight we do not insist that if we start off with only objects that are circular brass discs with a knob in the centre then the concatenation of two of them must produce a brass disc with a knob in the centre. Thus, more generally, we may fall back on the position that in making a measurement we are not necesarily counting the unit objects that comprise the measured object itself but counting those that would be needed to build up an object of the same magnitude.

However I not think that we are obliged to make even this small concession. It is necessary only if we take an unduly restricted view of what the unit elements that we are attempting to count must be. The initial objection rests on the view that the unit elements to be counted must be bits of the physical structure of the object which possesses the property being measured. This may be so in some cases, but as a general statement it is plainly false. There is no difficulty in thinking of the weight of a lump of brass as the sum of the weights of the unit sized pieces into which it can be cut. But the time period of a simple pendulum is never supposed to be the sum of time periods of a set of smaller pendulums which can be obtained by cutting up the larger one. Indeed given that the time period of a simple pendulum does not vary linearly with its length there seems to be no way even to contrive this. In this latter case the measurable entity is a time interval defined by the swing of the pendulum and what it may be decomposed into is smaller time intervals. So it is with the objects of Ellis's system. The measurable entity is an interval in space defined by the extremities of the object, and, according to the prescription for concatenation in this system, an interval 2 dinches long, whether defined by the extremities of a straight rod or a bent one, is composed of two intervals at right angles, each 1 dinch long.

I conclude therefore that the existence of the dinch scale does not threaten the idea that extensive measurement is counting.

CHAPTER FOUR

CLASSIFICATION of SCALES

4.1 Types of measurement scale

The account of numerical representations given in Chapter 2 was confined to a comparison of ordinal and extensive measurement and, correspondingly, of ordinal and ratio scales. The main purpose was to exhibit the foundations of extensive measurement and to show how these lead to the construction of a ratio scale. However there are other types of measurement and other types of scale. Roughly speaking, by the phrase "type of measurement" I mean the kind of procedures used to obtain the order C_1, C_2, C_3..... whereas by "type of scale" I refer to the mathematical features of the particular numerical assignment that is made once the order is set up. There are difference measurement, derived measurement, associative measurement, and there are linear interval scales, logarithmic interval scales, and so on. It seemed natural enough in the light of our analysis to suppose that a ratio scale is peculiarly appropriate to extensive measurement, though as we saw this supposition is contentious. In other cases links between types of measurement and types of scale are less clear cut. I think the method of description that we have applied to ordinal and extensive measurement can be useful as a basis for examining these other cases as well. In this chapter we shall widen the area of discussion, and give an account of a range of other sorts of measurement. Among other things I hope to throw further light on the issue of the conventionality of scales.

Various classification schemes have been proposed, for example by Campbell (1920), Stevens (1946), Coombs (1952). Ellis discusses and categorizes such schemes (1966, Ch.IV). Broadly speaking there are two approaches. Campbell's scheme highlights the procedures involved in constructing the scales, and is more a classification of types of measurement than of types of scale. Those of Stevens and Coombs on the other hand are based on the mathematical characteristics of the scales themselves. What I propose to do is to try out a different method of classification, the one to which I alluded in the introductory mateial of Chapter I. The method is suggested by our analysis of the calibration programme for weight on which we have so far based the discussion. It is a classification in terms of the kind of information that is incorporated in the construction of a scale. It is perhaps quite close to Campbell's, but I think it is sufficiently different in emphasis to be worth spelling out. I shall begin with an explanation of the scheme, and then follow with a survey and discussion of a variety of types of measurement.

As the account given in Section 2.1 was designed to show, the information on which a ratio scale is constructed comes from two sources. One is the results of a set of operational tests. These results determine the relation $\succsim$ on the set A which is the basis of the weak order. The other is knowledge of the structure of objects in A. It is necessary to know about how certain objects in the set are constituted from others. This kind of information about the objects is clearly different from the operational information, it relates to a different relation from $\succsim$, namely the relation that exists between two objects one of which is a constituent of the other. Without it we could not make the transition from the ordinal to the ratio scale. The fact that information about the two sorts of relation are incorporated in the numerical assignment is reflected in the fact that the assignment is governed by the two rules that were given in the Chapter 2:

I For any pair of objects, x and y, the corresponding numbers n(x) and n(y) are such that:

(a) if x is at a higher position in the order than y then n(x) > n(y);

(b) if x and y share the position then n(x) = n(y).

II For any body x that is a composite of separate bodies y and z the corresponding numbers n(x), n(y) and n(z) are such that:

$$n(x) = n(y) + n(z).$$

As we have seen, provided that the order is fully determined and a value has been chosen for one element, the second rule removes the arbitrariness of choice left open by the first and allows all other values to be fixed uniquely.

This is a general phenomenon of measurement. According to standard theory the existence of a weak order, based on some operational relation, is a characteristic of all measurable physical properties, extensive and otherwise. The account of the sub-programme which led to an ordinal scale for weight can be adapted to fit any other case, and it describes a logically prior step in the construction of every scale. Rule I applies for any scale. However if anything other than an ordinal scale is to be constructed we require information from elsewhere about other relations of objects in the order, together with a second rule for incorporating it into the scale. This additional information, and correspondingly the second rule that goes with it, can be of several different kinds, and this in turn leads to the variety of types of scale that are found in practice. The nature of the information determines the type of scale. This point is illustrated in the following survey where each measurement type is dealt with as follows. We first assume that a weak order is obtained on a set of objects by means of a binary operational relation which as before can be represented as a simple ordering of equivalence classes:

$$C_1, C_2, C_3, C_4, \ldots$$

We then ask: what extra information about these classes, either about their number, about other relations among them, or about their membership, can be used to construct a scale, and how does the nature of the information affect the characteristics of the scale?

We shall start by making further remarks about ordinal scales in which, as we have already seen, no extra information at all comes into play.

4.2 Ordinal Measurement

This type of measurement is based solely on information about the relation $\succeq$ gained from the operational tests. No other information is exploited; indeed there may be none available. In this situation there is no choice but to resort to an ordinal scale. One is Moh's scale of hardness in mineralogy. For this, solid substances are arranged in order according to the outcome of a scratching test. Substance x is harder than substance y if a sample of x will make a scratch on a sample of y. No further information is used.

One consequence of this lack of information is that the choice of a single reference standard is of little value. Suppose a particular object s from the ordered set is chosen as a standard and given the value n(s). The value n(x) for some other body x is then sufficient to convey whether x is greater than, equivalent to, or less than s in the property in question, but that is all. The only way in which things can be improved with such a scale is to choose as many standards as is practicable distributed conveniently along the order so that knowledge of the value n(x) then allows x to be located between two of them. This is done for example in the construction of a scale such as the Richter scale for the measurement of earthquake intensity. For this a set of 12 types of seismic event is selected, each type being specified in terms of certain recognizable effects. Each specification is intended to be sufficiently precise to define a standard. The set is then arranged in a series in order of increasing intensity (it is assumed that the means exists for doing this) and numbered from 1 to 12. Any actual event can then be evaluated by matching its observable effects against these standards. In the case where it is judged to fall between two adjacent members of the set of standards it can be assigned a fractional number between the two. In principle there is no significance attached to which intermediate number is chosen, except in so far as an observer makes a judgement on grounds that may not be well articulated that the event is closer in nature to one of the standards than to the other and so chooses a number in the appropriate half of the interval.

4.3 Distribution Measurement

The next type I wish to mention has not been categorized, so far as I know, under a commonly accepted name, but I think it is sufficiently distinct to merit separate recognition and I have coined the term "distribution measurement". It applies to cases where the distribution of objects among the classes C_1, C_2,... displays stable characteristics and where the scale number assigned to a particular object is then meant to represent the position of that object in the distribution. We can explain the point with a contrived trivial example, and then go on to give a more serious one.

Suppose that a scale is required to measure size of orbit for the planets of the solar system. The planets may be ordered according to some suitable relation. For example we could interpret $\succsim$ so that $x \succsim y$ holds if the orbit of planet x encompasses the orbit of planet y. We may then construct an ordinal scale with numbers arbitrarily chosen within the constraints of rule I. In this case the most natural choice would be the integers 1 to 9. This is because, in addition to knowing about the order, we know there are only nine in number. The result is that these numbers are more informative than any less systematically selected numbers would be. They do not tell us anything more about the magnitudes of the orbits, but they do convey extra information of another kind. We can deduce for example that the planet whose orbit size is 4 encompasses 3 other planets within its orbit. It is as if we adopted a second rule of assignment:

III For any object x, n(x) is to be the (cardinal) number of objects lower than x in the order.

This is possible by virtue of the fact that the calibration is performed on a set of objects which remains fixed. The scale is not intended for application to further objects to be added subsequently the set. This is a case where the distribution in the order is stable for the most elementary of reasons. The order consists of just nine, single membered, equivalence classes, and this distribution is not expected to change. Indeed if more planets were discovered the existing values would thereby cease to carry their extra content. The simple scale just described measures how each object stands with respect to a fixed distribution.

A more serious example is a scale for measurement of intelligence. For this a suitably large set of individuals is ordered according to scores in a test. In this example the set is not in general fixed, but the measurement procedure is based on the assumption, which is open to test, that the position of any individual in the order will be (approximately) the same for any randomly chosen set of the same size. That is to say the distribution of individuals throughout classes C_1, C_2,... etc. is assumed to be stable. The I.Q. is an index of how the individual stands in that distribution. The number 100, for example, indicates that he or she is at the centre of the order (with due modification for age variations if necessary). The calculation of I.Q. follows a more complex rule than that in (2a) but essentially it rests on counting the number of individuals lower (or higher) in the order. A difference in the values of I.Q. for two individuals does not indicate the amount by which one exceeds the other in intelligence, supposing there is any sense at all in the phrase "amount of intelligence"; it merely indicates the percentage of the population that falls between them both in performance in the tests.

4.4 Interval (or Difference) Measurement

Consider the problem of constructing a scale which measures the position of points on an axis, as for instance an axis of a Cartesian coordinate frame. These points are ordered by virtue of their position along the axis. Standard scales are based on the definition of equivalence of intervals along the axis. The definition can be given in operational terms. For example, the points a, b, c, d, etc. shown here are counted as being equidistant if a rigid rod whose extremities coincide with a and b, can be brought into coincidence with b and c, c and d, and so on, as the rod is moved along the axis.

Fig.4.1 . . . . . .
a b c d e f

These points may then be numbered off starting from one arbitrarily chosen as the origin. If a in the above figure is the origin, then the value 3 for d registers the number of equivalent intervals separating it from a. Thus the additional information that is brought to bear in making this kind of numerical assignment concerns the spacing of the elements, not simply the ordering. The term "spacing" here is not meant to imply the prior existence of some metric, but simply that there is some procedure, of the sort we have described, which reveals an appropriate binary relation between pairs of elements.

There is a deal more to be said about this method of construction, of course. It has the shortcoming that it leaves out of account the remaining intermediate points. As things stand it does not include any systematic procedure for assigning values to any of them. The situation can be improved by choosing a shorter rigid rod, and the shorter the rod the greater the number of points that come to be included, but points within these smaller intervals will still remain unevaluated. There is the theoretical possibility that space is quantized, and that a procedure may ultimately be devised - employing a quantum rod - for picking out pairs of adjacent discrete points. These pairs could then be regarded as defining successive equivalent intervals, all points could be numbered and, in some sense, an absolute scale would thereby be established. However in default of this situation's being realized, the limitation we have mentioned is unavoidable.

The calibration process we have described is entirely equivalent to fixing the scale by measuring the distance of a point from the origin. According to this way of looking at it the intervals, defined by pairs of points, are treated as entities to be included with a suitable collection of rods to make up a set on which a scale of length is constructed. The problem of the failure to assign values to points which are too closely spaced is translated into the problem of the limit of precision determined by the size of the elements in the unit class. This unit class will contain the smallest rods in our collection. There may well be smaller intervals on the line, but they cannot be incorporated in a fully determined order because there is no way of showing that these smaller intervals are equivalent to any other intervals. They cannot themselves be put into an appropriate equivalence class.

Thus constructing an interval scale on a set of points on an axis is equivalent to constructing an extensive scale on the intervals defined by them. Whichever description is adopted, the numerical values assigned to points on the axis are effectively arrived at by counting the number of unit intervals which separate the point from the origin.

4.5 Extensive Measurement

This category is mentioned again for the sake of completing the list and there is nothing substantial to add that did not emerge in the account of the calibration programme given in an earlier section. The main point to reiterate is that in this case the additional information needed to determine the numerical representation comes from knowledge of the composition of individual objects in the set. This is the first type of scale in this list which draws upon information about the structure of the objects themselves.

4.6 Associative Measurement

An associative scale is another type in which the added information comes from the nature or the structure of the objects themselves. The term "associative" has been used, e.g. by Ellis (1966, Chap. IV), to describe scales in which values are assigned to objects with respect to one property on the basis of values already assigned to them with respect to a second property that is known to vary in association with the first. The method presupposes that a scale for the second property has previously been constructed by a procedure that is independent of any measurement of the first property. (Use of the term is not meant to refer to associative properties of any mathematical operations involved in the construction of the scale. Perhaps "associated measurement" would be better.) A prime example of this type is a so-called empirical scale of temperature. The entities which make up the membership of the classes C_1, C_2,... are temperature states of different bodies or physical systems. In general, for any given system there is one member of each class that is a state of that system. The usual procedure for constructing an empirical scale involves choosing one type of system as the thermometric system, any particular instance of which can therefore be employed as a thermometer. Some property of the thermometric system, for which a scale of measurement already exists is chosen as the thermometric property. Examples are the length of a column of mercury in a mercury-in-glass thermometer, the pressure of a gas in a constant volume gas thermometer, the resistance of a platinum wire in a resistance themometer. Details are to be found in introductory texts, e.g. Zemansky and Dittman (1981, Section 1-7 *et seq.*).

The necessary requirements for a thermometric property are that it should have a unique value for a given temperature state, i.e. just one value per equivalence class, and that the values for different classes should be distinct and should increase monotonically along the order. These values may be referred to as the thermometric values of the equivalence classes. Values are then assigned to temperature states in accordance with a rule of the general form:

IV If x is a member of the equivalence class whose thermometric value is p then $n(x) = F(p)$.

Here F is some well behaved strictly increasing function. The choice of the function F is conventional. In certain important standard scales it is of the form:

$$F(p) = \text{constant} \times p$$

e.g. the constant volume gas scale where p stands for gas pressure. If a function of this form is adopted it follows that the thermometric property varies linearly with the temperature as it is defined on this scale. The pressure of a constant volume of the gas specified in the construction of a gas thermometer increases linearly with temperature as measured on the gas thermometer scale. It is clear that this linearity is purely a consequence of the definition, it is not a pre-existing fact providing a reason for choosing the thermometric property in question in the first place. To this extent then there is a large conventional element in the construction of an empirical scale. There is a choice of which thermometric system; there is a choice among the properties of that system of which thermometric property; there is a choice of the function F.

The general basis of an associative scale appears to be this. Once the order of classes C_1, C_2,... has been established (on the outcome of operational tests) it is noticed that there are various sets of numbers already attached to the classes. These are numbers which have previously been assigned to certain members in connection with some wholly separate purposes. An associative scale is defined by choosing one of these sets. It is a little like using a roll of cloakroom tickets to construct an associative scale of length. We notice that the tickets have numbers printed serially on them, and, irrespective of whether or not all tickets are of equal length, we use the roll as a tape measure taking the ticket numbers, or the values of some function of them, as scale values.

Before we move on, it should be pointed out that in addition to associative measurement, interval measurement of temperature is also possible and the two must be distinguished. Interval measurement is achieved by applying the general procedure of Section 4.4 to an ordered set of temperature states. There are various possible ways of defining equivalent temperature intervals. One is to count two intervals as equivalent if the same quantity of energy is required to raise a given mass of some substance through each interval. For example if 1 calorie of energy is absorbed by 1 gramme of water its temperature rises by an amount we may stipulate to be 1 degree. If it is then allowed to absorb another calorie of energy there is a further rise and this, by definition, is another 1 degree change. In this way 1 degree intervals are stepped out along the temperature axis just as 1 metre intervals are stepped out along a straight line. Another very important possibility is to define equivalent intervals in terms of the operation of an ideal Carnot heat engine. Such an engine operates in thermal contact with two heat reservoirs at different temperatures. This is represented schematically in the following Fig.4.2 where reservoir 2 is at a higher temperature than reservoir 1.

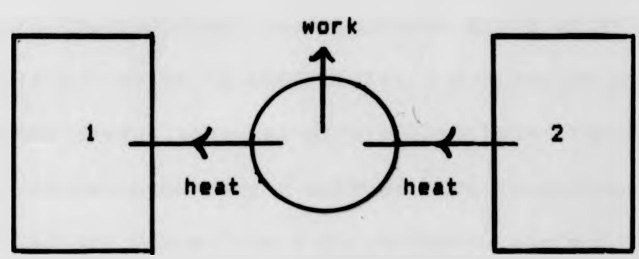


Fig.4.2

The point is that the efficiency of the Carnot engine depends solely on the reservoir temperatures, increasing as the difference between them increases. Two such reservoirs clearly define a temperature interval. Thus two intervals may be defined as being equivalent if a Carnot engine is equally efficient when working across either of them; and if the series of reservoirs shown in the Fig.4.3 (next page) are at temperatures such that the engine works with the same efficiency between any adjacent pair then they constitute a series of equivalent steps on the temperature axis.

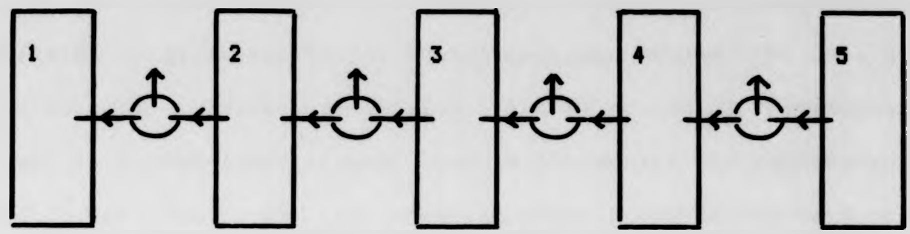


Fig.4.3

The modern Thermodynamic (or Absolute) Scale is effectively based on this principle and so it is essentially a difference measurement scale. The matter is complicated somewhat by the fact that, for mainly historical reasons, the scale has undergone a mathematical transformation. On the standard scale, values range from 0 to infinity, where 0 is assigned to a state referred to as "the absolute zero of temperature". On this scale, the efficiency, e, of a Carnot engine is given by the formula:

$$e = 1 - (T_1/T_2),$$

where T_2 and T_1 are the temperatures of the higher and lower reservoirs respectively. This gives the same value of e for the same values of the ratio (T_2/T_1) but not for the same values of temperature difference $(T_2 - T_1)$. However consider the effect of transforming to a scale on which the new temperature value t is related to the old by:

$$t = \log T$$

On this scale we have for the efficiency of the Carnot engine:

$$e = 1 - \exp(t_1 - t_2)$$

which now does take the same value for equal intervals in t. On this new scale values range from minus infinity to plus infinity, with the absolute zero now banished to the former extremity. It might be thought that this constitutes an objection to the transformation. However the absolute zero of the present scale is unattainable - the Third Law of Thermodynamics says so. It is not like the empty pan in the scheme for measuring weight. Indeed in very low temperature physics, where temperatures very close to zero (on the standard scale) are attained, say below one thousandth of a degree, it is common practice to transform the scale in the way I have suggested in order to stretch out the range.

In that it has called for lengthy explanation this account of the measurement of temperature is in danger of being a digression. But a number of points have emerged that are directly relevant to the main line of the discussion. We have seen that temperature states are amenable to both associative and interval measurement. This contrasts with the case of points on an axis where only the latter type is possible since points, unlike temperature states, do not possess individuating characteristics on which to construct an associative scale. In the history of temperature measurement associative measurement came first, largely I suppose because the type of information it calls for was more accessible. It is a major achievement of modern thermodynamics to have discovered the significance of interval measurement as well and to have brought about a wonderfully elegeant assimilation of the results of both methods. The form of the standard scale reflects the way in which this has been done.

However, accidentally, because of the particular assignment of the zero on the scale, it tends to convey a spurious sense of a ratio scale. It tends to suggest that there is something associated with temperatures of, say, $100^{\circ}K$ and $200^{\circ}K$ such that twice as much of it goes with the second as with the first. But so far as we know there is nothing answering this description that can be defined in terms of physical operations. There is nothing additive about temperature values because there is no generally applicable concatenation operation whereby two systems at equal temperature T, say, can be combined to give a system with a temperature greater than T. By contrast, it is possible to combine temperature intervals in a suitable way, and so these have additive properties of a precisely similar kind to those of distance intervals along an axis in space. This is a case where the distinction between, on the one hand, the type of measurement process which provides the foundation for the scale, and, on the other, the type of numerical scale that is adopted, is quite clear. The example lends support to the argument which was developed in Section 2.3 of Chapter 2 that, although from a mathematical point of view we have the freedom to transform scales in more or less any way we choose, the type of measurement used to construct the scale in the first place will depend upon the existence of certain nonconventional empirical attributes which in turn make one type of scale more appropriate than another.

4.7 Derived Measurement

Derived measurement is similar to associative measurement in that the additional information it draws upon relates to measures previously established for other purposes on members of the classes C_1, C_2,.... On the other hand it differs in one important respect. In associative measurement the measurable property that is exploited generally belongs only to selected members of each class. In the case of temperature, for example, a given equivalence class contains temperature states of many different types of system, homogeneous samples of gases, liquids, and solids, resistors, thermocouple junctions, radiation-filled cavities and indefinitely many more. A particular thermometric property belongs to only a few of these, and among this few the sets that share a common value of the property will be smaller still. Once the choice is made the value that truly belongs to the thermometric system is attached to the others by association. In derived measurement on the other hand, the measurable properties that are exploited, and the particular value they take, are common to all members of a particular equivalence class.

A good example is a derived scale for density. It is assumed that we have an operational test for comparing objects with respect to density, involving floating or sinking in liquids, say, and that through this an order in the form of a sequence of equivalence classes is obtained in the usual way. The members of the classes are bodies of various substances. There is no feasible concatenation operation for density measurement, no generally applicable operation whereby for instance two samples of the same substance are combined to give a sample of greater density. There is no question therefore of constructing a scale by extensive measurement. However it can be assumed that scales for the extensive properties mass and volume are already established. It is noticed that (i) for any given equivalence class the ratio of mass to volume, M/V, is the same for every member of the class, and (ii) the value of this ratio increases monotonically along the order. This much is an empirical discovery. The usual convention is then to adopt the values of this ratio as the density values. One useful feature of this sort of analysis is that it makes it easy to disentangle the empirical content from the conventional part of an equation such as:

$$D = M/V$$

This equation may be expanded to read:

$$(1) \quad (\exists f)[D = f(M/V)], \quad \text{and} \quad (2) \quad f(y) = y$$

(1) is the mathematical expression of the discoveries just listed in (i) and (ii), and it is a law of physics. (2) on the other hand expresses a conventional decision; in principle any other well behaved function would do. Thus although the equation $D = M/V$ is usually regarded as defining density, it is clear that it does so only in so far as defines the scale values. The independent existence of the property of density is already established once the set of objects has been ordered on the outcome of the operational tests. The subsequent finding that the ratio M/V has some significance in the order is a further empirical discovery.

The division between empirical content and conventional content does not always fall in the same place. Consider as a final example a derived scale for momentum. We suppose that here the set to be ordered is a set of states of bodies, where a state is specified by the speed of the body. We may devise an ideal operational test for comparing momentum states. We arrange for two bodies p and q in specified states x and y to approach each other as shown here and they are allowed to collide and coalesce.

Fig.4.4

If the resulting composite body is at rest or continues to move to the right then we have $x \succsim y$. In this case a suitable concatenation operation is also possible. For this we allow p and q to move in the same direction along the same axis with the faster moving one catching up from behind.

Fig.4.5

They again collide and coalesce, and the resulting composite body is deemed to be in state $x \circ y$. This body in this state is available to be compared with some other, and so on. On the basis of these two procedures we may obtain a fully determined order of momentum states C_1, $C_2, \ldots,$ in a manner entirely similar to that for obtaining the order for weights described in the calibration programme of Section 2.1.2. This is the foundation for a fundamental extensive scale of momentum. However, on an examination of the members of each C_k it is found that (i) for any given class the product of mass and speed $M \times v$ is the same for every member of the class, and that (ii) the value of this product not merely increases monotonically along the order (which was the case in the example of density above) but is proportional to the value of momentum assigned to that class by the fundamental measurement procedure. Thus with regard to the familiar equation:

$$P = M \times v$$

of mechanics, (where "P" stands for momentum), the division of empirical content and conventional content is best represented by rewriting it as:

$$(1) \quad (\exists k)[P = k \times M \times v] \quad \text{and} \quad (2) \quad k = 1.$$

This completes the survey of important types of scale.

CHAPTER FIVE

MEASUREMENT OPERATIONS and NONCOMPARABILITY

5.1 Axioms for a Weak Order

This chapter is devoted to an investigation of a problem concerning the definition of a weak order. The definition has already been stated in Chapter 1 (Def.1.2), and we restate it here, now renumbered as Def.5.1.

Def.5.1 Let A be a set and $\succsim$ be a binary relation on A, i.e. $\succsim$ is a subset of AxA. The relational structure $\langle A, \succsim\rangle$ is a *weak order* iff for all $x,y,z \in A$, the following axioms are satisfied:

1 $x \succsim y \;\vee\; y \succsim x$.
2 $(x \succsim y \;\&\; y \succsim z) \rightarrow x \succsim z$.

5.2 Direct Comparison Operations

The problem I wish to discuss is related to the interpretation of the relational predicate $\succsim$ of this formal definition, or more particularly with an interpretation of a certain type. In standard accounts it is usual to give an operational interpretation, one that determines the truth value of $x \succsim y$ according to the outcome of an operational test. An example is the following possible interpretation of $\succsim$ in the application of the structure $\langle A, \succsim\rangle$ to weight measurement.

I.5.1 The expression $x \succsim y$ holds iff, if objects x and y are placed on opposite pans of a balance, the balance remains in equilibrium or else the pan containing object x descends.

A crucial feature of I.5.1 is that the operation it describes is a *direct* comparison of x and y in the course of which the two objects mutually interact in a specified experimental arrangement. This is a feature common to interpretations of $\succsim$ standardly assumed in other examples. In the case of length we may have:

I.5.2 The expression $x \succsim y$ holds iff, if rods x and y are laid side by side with one end of each in coincidence, the other ends coincide or else x extends beyond y.

In the case of temperature the interpretation may be:

I.5.3 The expression $x \succsim y$ holds iff, if a system x [a specified body in a specified state] is in thermal contact with a system y, the two are in thermal equilibrium, or else heat flows from x to y.

These and many others involve a direct mutual interaction of the pair of objects x, y. It is this feature that gives rise to the problem I wish to discuss.

5.3 Conditional Interpretations

The problem is that as they stand these interpretations do not, in general, apply as intended to all pairs of objects in the set. Consider I.5.1 and suppose that x and y are identical. x and y cannot be compared directly in that case since the configuration required for the operational test is not realizable. It is impossible to place the same object on both pans of a balance at the same time; a single physical object cannot be in two different places at once. Similar considerations apply for identical elements in other examples. The problem may not be thought to be equally serious in them all. It is clear enough that we cannot get the same object on two pans of a balance at the same time, but perhaps the idea of laying down a rod side by side with itself is not so difficult. It could be held that what we really do when we compare two objects for length is arrange for two lines - usually edges - to coincide. This coincidence obtains as a matter of necessity for a single object, and the expression $x \succsim x$ well signifies this. If this argument is accepted then we may simply omit this example, but the others remain to be considered. The problem for them is this. The axioms require $\succsim$ to be reflexive ($x \succsim x$ follows directly from Axiom 1 for $x = y$). An acceptable model must provide an interpretation that makes the expression $x \succsim x$ true. However it is not clear that the ones we are discussing meet that requirement.

At first sight it looks as if the answer lies in the fact that I.5.1 and the other examples are conditional statements. They are rules for determining whether $x \succsim y$ holds, expressed in the form: if such and such an operation is performed with the pair x,y then such and such an outcome is obtained, or in shorthand:

$$\text{if } P[x,y] \text{ then } Q[x,y].$$

We may refer to such an interpretation as a <u>conditional interpretation</u>. The fact that $P[x,x]$ is not realizable, it might well be thought, is not a problem because we may treat the interpretation as a material conditional. For the case of $x = y$ the antecedent is false and $x \succsim x$ comes out true. However a simple argument will show that this is not a satisfactory solution. Though it is alluring and certainly disposes of the problem of reflexivity, further probing reveals that we cannot let it rest there.

Suppose that we are investigating properties of a relation (which, using obviously suggestive notation, we can denote by "≻") that has the following interpretation:

I.5.4 The expression x ≻ y holds iff, if objects x and y are placed on opposite pans of a balance, the pan containing object x descends.

This is obtained from I.5.1 by removing the reference to equilibrium. Is this relation reflexive also? If we appeal to the fact that I.5.4 is a conditional statement the answer must again be yes. Intuitively, however, we are inclined to give the opposite answer for this example. We choose to assume that ≻ is irreflexive, because this assumption leads to a more complete correspondence between ≻ and the numerical relation >. But why should we be entitled to ignore the truth functional nature of I.5.4 and make an alternative decision based on what appear to be purely pragmatic grounds, namely the need to satisfy requirements of the structure we aim to build on the relation? If this is legitimate then surely I.5.4, at least when treated as a material conditional, is inadequate. In that case there is doubt about the role of I.5.1 in relation to ≿. For the same considerations apply. The decision that ≿ is reflexive must similarly be determined by how the assumption fits in with the total pattern of relations, and the fact that it is also in accord with the truth functional reading of I.5.1 is gratuitous. Unless there are reasons for distinguishing the two that are independent of the formal requirements, a reading of I.5.1 and I.5.4 as material conditionals is acceptable for both of them or else for neither of them.

Similar considerations apply if we attempt to solve the problem by replacing I.5.1 with a subjunctive conditional, i.e.:

I.5.1′ The expression $x \succsim y$ holds iff, if objects x and y were placed on opposite pans of a balance, the balance would remain in equilibrium or else the pan containing object x would descend.

This renders $\succsim$ reflexive, but a corresponding replacement for I.5.4 would in the same way render $\succ$ reflexive. If anything the situation is worse. Under I.5.1′ the relation $\succsim$ now holds between any pair of objects that simply happen not to have been compared with one another, whether because of logical impossibility or simply because of physical impracticability. There is the well known problem of counterfactual cases.

If this difficulty were confined to the question of reflexivity it would perhaps be seen as no more than a minor point of irritation, and hardly worth pursuing. But, as we now go on to demonstrate, it is only part of a more general and more substantial problem.

5.4 Noncomparable Pairs

There are many possible applications where the set A will contain as distinct elements different states or different features of the *same physical object* that cannot therefore be compared by a direct method. Consider the example of temperature. Here the members of A are specified states of bodies or systems, and under I.5.3 above $x \succsim y$ is to be understood in terms of direct thermal contact between bodies in specified states. However if the two members of A concerned happen to belong to the same body this cannot be done. A block of copper at 50°C can be placed in thermal contact with a different block of copper at 20°C but not with the same block of copper at 20°C.

Temperature is not an isolated example. The same applies to any measurable property that takes more than one value for the same physical object. This happens for most if not all physical properties: temperature, speed, volume, magnetic dipole moment, and many more. Even the mass of an object varies, according to modern relativistic theories. The situation we have described is quite general. Thus the more usual situation is that the elements of the set A are not objects but *states* of objects and in such cases the specification of an element both refers to a particular object and describes the state that the object is in. Some pairs of states are associated with the same physical body and are therefore restricted in the way that we have described.

In mechanics for example a body p moving with a speed v_1 may be regarded as one element, the same body p moving with speed v_2 a different element. A different body q moving with speed v_1, and then later with v_2 gives a third and a fourth, and so on. In thermodynamics a fixed sample of a gas can be found in various states, different states corresponding with different pairs of values of the volume and pressure of the sample. Such states, belonging to a whole set of samples of gas, may be taken as the elements of A. In magnetism, we may wish to count as distinct members of A different magnetic states of the same piece of magnetic material. It is clear that in all these examples there are pairs which cannot be compared by a direct method. Consider the first example and recall the description we gave above in Section 4.7 of a possible method of ordering a set of bodies, moving at various speeds, according to momentum. The direct test, depicted in Fig.4.4, requires two (necessarily distinct) bodies to approach each other at specified speeds so that they collide and coalesce. But this method is useless when we come to compare successive momentum states of the same body. It is impossible for a body moving at one speed to interact in this way with the same body moving at a different speed.

Further examples can be cited indefinitely, and the situation for all of them is that there are indefinitely many pairs of states for which the operation P cannot be performed. We shall refer to any pair x,y of members of as *comparable* or as *noncomparable* (relative to the operation P) according to whether P[x,y] is realizable or not.

It is readily seen that under a conditional interpretation the existence of noncomparable pairs leads to difficulties. For any such pair the antecedent is false, and both $x \succsim y$ and $y \succsim x$ hold. This gives the disconcerting result that the relation $\sim$ holds among all the states of a single body. It is quickly seen that this then leads to failure of the axioms for a weak order. For suppose there is a third object z that *is* comparable with both x and y such that $x \succsim z$ holds but not $y \succsim z$. This violates the condition of transitivity (i.e. Ax.2 of Def.5.1).

For good measure we may mention corresponding arguments which arise with respect to the structure $\langle A, \succsim, \circ\rangle$. Here the concatenation operation produces further examples of pairs for which these restrictions occur, pairs such as $x \circ y$ and x, or $x \circ y$ and $x \circ z$, which have some part in common. Problems similar to those described above arise in relation to expressions such as $x \circ y \succsim x$ and $x \circ y \succsim x \circ z$. The situations to which they refer, on a straightforward reading of an interpretation such as I.5.1, are not realizable. If the interpretation is treated truth functionally the consequences are as bad as in the last example. For any $x,y,z \in A$ we have the result:

$$x \circ y \succsim x \circ z \;\; \& \;\; x \circ z \succsim x \circ y, \;\; \text{i.e. } x \circ y \sim x \circ z.$$

Any composite object is held to be equivalent to every other composite object that has a constituent in common with it, and it can be shown by an argument similar to that given in the last paragraph that the weak order axioms fail.

We conclude that in general, under conditional interpretations of the sort shown in I.5.1 to I.5.3, the structure $\langle A, \succsim\rangle$ is not a weak order, and falls short as a representation of the physical property involved. Either the interpretation or the structure must be modified or replaced and I wish to examine ways in which this can be done. The matter is of some importance from a number of points of view.

For one thing, if it is accepted that any measurable property is a characteristic not of objects in isolation but of their relation to other objects of a similar kind, successful analysis of measurement will depend on our having an accurate picture of the *pattern* of relations in question. What has been said so far shows that the pattern is more complex than that of a weak order based on a simple direct operational relation.

Our problem also bears on the issue of the nature of the relation between measurement operations and measurable properties. According to an operationist view for example, measurable properties are defined by the operations used to measure them. Measurement is a process of invention. Dingle, for instance, refuses to accept that quantities exist prior to, and independent of, measurement operations and insists that "we should begin with the operation and its result, and then if we wish to speak of a property (which I do not think we shall do) define it in terms of that", (1950). On an anti-operationist view on the other hand measurement is more a process of discovery. Quantifiable properties like weight do exist independently of the means of measuring them, specific operations are employed to reveal these properties, and we are free to judge whether or not any proposed operation is adequate for the purpose. It is clearly important for this debate to establish whether or not it is possible to formulate a satisfactory interpretation in purely operational terms.

5.5 Operations with Indirect Comparison

On the face of it there are two alternative approaches to altering I.5.1 and the other examples so as to avoid the problem of noncomparable pairs. The more radical of the two involves changing P. We could eschew operations involving direct comparisons of the sort that I.5.1 and the others call for, and adopt instead operations with respect to which all pairs are comparable. Whether or not there are operations for which this condition is satisfied is a matter for investigation. The most promising type of candidate is an operation in which two objects are compared by comparing the outcomes when one is substituted for the other in some experimental arrangement that otherwise remains fixed. For example two objects x and y may be compared for temperature by putting each of them in succession in thermal contact with a suitably chosen third object z and noting the direction of heat flow, if any, in each case. The force of the qualifier "suitably chosen" is that for the test to be effective z must be either equal in temperature to one of x, y (or to both if x and y themselves are equal) or else lie between them in temperature. The appropriate interpretation for $x \succsim y$ could be formulated as follows:

I.5.5 The expression $x \succsim y$ holds iff there is a system z such that if system x is in thermal contact with z the two are in thermal equilibrium or else heat flows from x to z, and if system y is in thermal contact with z the two are in thermal equilibrium or else heat flows from z to y.

The possibility of performing this operation is independent of whether x and y are states of the same or of different bodies. It does depend on the availability of suitable intermediary systems, however. For each x,y there must be an appropriate z, or once again there will be noncomparable pairs. Further, if each z is a member of the set, there have to be intermediary systems to allow comparison of x with z and y with z, and so on. Thus I.5.5 imposes fairly strong existence conditions. There is no difficulty in meeting these in most instances, but the implications for the character of the interpretation should be noted. It is clear that I.5.5 gives a sufficient *rule* for determining if $x \succsim y$ holds. But it hardly gives an acceptable *definition* of $x \succsim y$. The idea that "x is hotter than y" includes reference to the existence of a third object is not part of our usual understanding of the phrase. Suppose that at some moment there is no object in the world with a temperature between say 1°C and 4°C, except only for one object x at 2°C, and another y at 3°C. This is exceedingly improbable but not logically absurd. At that moment, if I.5.5 were the definition, y could not be said to be hotter than x, but it could subsequently be rendered hotter than x by the action of cooling a third object z from 5°C to 2.5°C. It is highly counterintuitive to suppose that x's being hotter than y depends upon the contingent matter of whether or not there is a third body equal in temperature to one of them or intermediate between the two. Despite this particular difficulty there is rather more to be said for this particular approach and we shall be discussing some further aspects of it in Chapter 6. However in the present chapter I am interested in a second more conservative approach to the problem, in which we attempt to stay as close as possible to the direct operation as the basis for the structure.

5.6 Construction of Ordering Relations from Operational Relations

First a point of nomenclature. We shall refer to the relation that is determined by comparing pairs directly, each in a single operation, as the *direct relation* or *D-relation*, and hereafter we denote it by the symbol "$\succsim_D$".

We now explain the general strategy for solving the problem as it arises in any given application. It falls into two parts.

The first step is to replace the conditional interpretation of $\succsim_D$ by what we may refer to as a *categorical* interpretation. This is one that has the form:

$$R[x,y] \text{ and if } P[x,y] \text{ then } Q[x,y].$$

The added clause R[x,y] is some condition guaranteeing that x,y is a comparable pair. It could simply be the (modal) statement that P[x,y] is possible, or alternatively it could express some further relation which when it holds between x and y ensures that P[x,y] is possible. An obvious possibility here is the relation (between states) of being states of physically distinct bodies. Either way R is intended to deliver the set of pairs that are comparable relative to P. We may, for example, replace I.5.3 by:

I.5.6 The expression $x \succsim_D y$ holds iff x and y are states of physically distinct systems and, if the system of state x is in thermal contact with the system of state y, the two are in thermal equilibrium or else heat flows from the system of state x to that of state y.

The adoption of a categorical interpretation amounts to a decision that $x \succsim_D y$ does not hold for any noncomparable pair x,y (by virtue of the fact that R[x,y] is not satisfied). $\succsim_D$ is a subset of the set of comparable pairs in A. Thus $\succsim_D$ is not connected on A. In particular if R[x,x] is not satisfied, as in this example, then $\succsim_D$ is also irreflexive. Clearly now $\succsim_D$ does not obey the axioms of Def.5.1, and $\langle A, \succsim_D\rangle$ is not a weak order. This is an instance of *essential* failure of the axioms.

In the second part of the procedure we define, in terms of $\succsim_D$ a second relation for which we retain the unsubscripted symbol "$\succsim$". $\succsim$ is intended to be the ordering relation, adequate for representing the measurable property involved. We shall refer to it as the *order relation*, or *O-relation*. An adequate O-relation will satisfy two conditions. The first is that for any pair for which $x \succsim_D y$ holds $x \succsim y$ must hold also. That is it must apply to any pair that were directly related in the first place. The second is that $\succsim$ must be connected and transitive, i.e. $\langle A, \succsim\rangle$ must be a weak order.

There are examples in the literature of structures in which the primitive relation is not connected, and in which it is necessary to define another relation to provide the order. One is due to Roberts and Luce (1968) which is based on an axiomatization of thermodynamics given by Giles (1964). They include in their structure an axiom of *conditional connectedness*:

$$(x \succsim_D y \;\&\; x \succsim_D z) \rightarrow (y \succsim_D z \vee z \succsim_D y)$$

This clearly does not hold for our D-relation as given by I.5.6. Under that interpretation this axiom fails for the case where y and z are states of the same system. In their scheme $\succsim_D$ has properties which, as it were, are the inverse of those it has in ours, in that theirs connects states of a *single system*, and fails to connect those of separate systems. An example for them is the relation which exists between two states of a single system when the first can be reached from the second by allowing the system to absorb heat. (This is different from that given by I.5.6. That a body gets hotter by absorbing heat is logically independent of the fact that heat flows spontaneously from a hotter to a colder body.) Thus the solution in Luce and Roberts depends upon a quite different choice of primitive relation from ours. Before I say more about their type of scheme I wish to show that we can define an adequate structure based on relations of the type in I.5.6. So far as I am aware, this particular problem has not received any attention in the literature.

5.7 Semi-Connected Orders

The problem is to define a relation $\succsim$ in terms of $\succsim_D$ and to give a set of axioms on $\succsim_D$ which ensure that the structure $\langle A, \succsim \rangle$ is a weak order. We shall continue to refer to the example of I.5.6 and to use the case of temperature as illustration.

$\succsim_D$ connects states of distinct bodies. To deal with pairs of states belonging to the same body we define, in terms of $\succsim_D$, a further relation $\succsim_I$ (which may be read as $\succsim_{Indirect}$) as follows:

Def.5.2 for all $x,y,w \in A$
$$x \succsim_I y \equiv (w)((w \succsim_D x \rightarrow w \succsim_D y) \ \& \ (y \succsim_D w \rightarrow x \succsim_D w))$$

This says that $x \succsim_I y$ holds when if anything bears the relation $\succsim_D$ to x it also bears it to y, and if y bears $\succsim_D$ to anything so does x. It is based on the principle of comparing two objects indirectly by means of a direct comparison with other objects in the manner described in Section 5.4. In this respect therefore the intuitive sense of $x \succsim_I y$ is akin to that given by I.5.5. There is the important difference however that Def.5.2 involves universal, not existential, quantification. The assertion that two objects stand in the relation $\succsim_I$ does not imply the existence of further objects.

If $\succsim_D$ is interpreted in terms of spontaneous heat flow from one object to another as in I.5.6 then $x \succsim_I y$ can be recognized intuitively as signifying that x is at least as hot as y. However strictly x is as hot as y in a different sense from that denoted by the relation $\succsim_D$. We might indicate this by coining two terms: "hot_D" (to suggest "direct") for use in connection with $\succsim_D$ and similarly "hot$_I$" for $\succsim_I$. Then x is said to be as hot$_I$ as y if anything as hot$_D$ as x is as hot$_D$ as y; and if anything no hotter$_D$ than y is no hotter$_D$ than x.

This new relation $\succsim_I$ has certain useful properties. It follows straightforwardly from Def.5.2 that it is both reflexive and transitive. (The proof of the latter is in fact spelt out below when it is needed as part of the proof of Theorem 5.1). Notice that these properties are independent of the interpretation of $\succsim_D$; they are part of the logical character of $\succsim_I$ itself.

However given the interpretation of $\succsim_D$ that underlies our present example $\succsim_I$ is certainly not connected. It applies only between states of the same body, and does not connect states of distinct bodies. To see this suppose that x and y are states of distinct bodies A and B respectively. We may assume that there is another state of B that is hotter - hotter$_D$ - than x. Taking this as an instance of w the first clause in the definition of $\succsim_I$ gives w $\succsim_D$ x but not w $\succsim_D$ y and hence not x $\succsim_I$ y. Thus all pairs that are comparable relative to $\succsim_D$ are noncomparable relative to $\succsim_I$ and vice versa. x $\succsim_D$ y and x $\succsim_I$ y are mutually exclusive.

There is just one detail in which the statement in the final sentence of the last paragraph may need to be amended. This relates to the question of the reflexivity of $\succsim_D$. We recall the comment, made in the initial discussion of the problem of conditional interpretations in

Section 5.3, to the effect that in some applications we might be perfectly happy to count $\succsim_D$ as reflexive. The example suggested was that of comparison of rods with respect to length under I.5.2. where the idea of comparing a rod with itself may be acceptable. In such examples if x = y the pair x, y is comparable with respect both to $\succsim_D$ and to $\succsim_I$. Now it so happens, as is easily checked, that Def.5.2 works in the intended way whether $\succsim_D$ is reflexive or irreflexive, and the analysis that we are about to carry out goes through in either case. We are free to settle our intuitive consciences on the matter how we will without affecting the results. This is perhaps little more than a technical curiosity but the point does crop up again in the discussion of semiorders in the next chapter and for that reason it has been worth pausing to mention it here.

However whichever way the question is settled the fact remains that the indirect relation $\succsim_I$ is no more connected on the set A than is the original direct relation. There is no question therefore of replacing one by the other to achieve connectedness, but the fact that together they do appear to cover all possible pairs prompts us to combine $\succsim_D$ and $\succsim_I$ to define a connected relation $\succsim$:

Def.5.3 for all $x,y,w \in A$,
$x \succsim y \equiv x \succsim_D y \vee x \succsim_I y$.

It is trivial that $\succsim$ satisfies the condition:

$$\text{for all } x,y,w \in A,\ x \succsim_D y \rightarrow x \succsim y,$$

which is one of the two conditions mentioned at the end of Section 5.6 for $\succsim$ to be acceptable as the O-relation.

Once an appropriate definition of $\succsim$ is obtained the next task is to state axioms on $\succsim_D$, that is, to define a structure $\langle A, \succsim_D \rangle$, such that $\succsim$ gives a weak ordering of A. There is of course a trivial solution, which is simply to state that $\succsim$ (as defined on $\succsim_D$) is connected and transitive. There is nothing in principle wrong with this. The work has been done, as it were, in devising Def.5.3, and indeed it has been carried out with an eye to satisfying the first of these conditions. The assumption that $\succsim$ is also transitive is open to test. However a subsidiary aim is to find axioms which are as simple and as intuitively clear as possible, and given its complex definition, axioms to the effect that $\succsim$ is connected and transitive hardly meet this aim. The axioms of the following structure - which I call a *semi-connected order* - seem reasonably clear, and, as we prove below, they are sufficient for establishing a weak order.

Def.5.4 Let A be a set and $\succsim_D$ a binary relation on A. If $\succsim$ is defined on $\succsim_D$ as in Def.5.3 the structure $\langle A, \succsim_D \rangle$ is a *semi-connected order* iff, for all $x, y, z \in A$, the following axioms hold:

1 $\neg(x \succsim_D y \vee y \succsim_D x) \;\&\; \neg(y \succsim_D z \vee z \succsim_D y) \rightarrow \neg(x \succsim_D z \vee z \succsim_D x)$,

2 $(x \succsim_D y \;\&\; y \succsim_D z) \rightarrow (x \succsim_D z \vee x \succsim_I z)$ [$\equiv x \succsim z$].

Axiom 1 is not a connectedness condition. Rather it expresses the transitivity of nonconnectedness. It is of some interest to note that the truth of this axiom is not settled by operational tests. In the examples we have been discussing, where it is states of single bodies or systems that are not connectable by $\succsim_D$, the Axiom 1 holds as a matter of physical necessity. The antecedent of the axiom holds if, and only if, x, y and z are all states of one body, in which case the consequent holds.

It is Axiom 2 that carries the burden of the operational tests. It expresses a kind of generalized transitivity of $\succsim_D$. Whether or not it is satisfied in any given example is open to experimental test. In some cases the question can be settled quite simply by appeal to commonplace knowledge, as with the commonly understood behaviour of a balance, for example. However in other cases the question may call for more difficult analysis of the empirical background. In the case of temperature for example the axioms must be interpreted in the light of the laws of thermodynamics. There is some question about the precise role of these laws in the foundation of a temperature order. In particular there is some discussion as to whether or not the Zeroth Law is independent of the Second Law, and if it is, whether or not the Zeroth Law is needed for the definition of temperature. (See, for example, Luckhardt and Kessler, 1971; Home, 1977; Erlich, 1981.) There is good reason to expect the result we have obtained here to throw some light on this problem; it would provide an interesting application of the analysis.

We now state and prove the following theorem.

Th.5.1 If the structure $\langle A, \succsim_D \rangle$ is a semi-connected order and if $\succsim$ is defined on $\succsim_D$ as in Def.5.3 the structure $\langle A, \succsim \rangle$ is a weak order

Proof: Connectedness.

Assume:

(a) $\neg x \succsim y$,

i.e. $\neg x \succsim_D y \;\&\; \neg x \succsim_I y$.

$\neg x \succsim_I y$ together with Def.5.2 gives:

(b) $((\exists a)(a \succsim_D x \;\&\; \neg a \succsim_D y) \vee (\exists b)(y \succsim_D b \;\&\; \neg y \succsim_D b))$,

and hence (a) gives:

(c) $\neg x \succsim_D y \;\&\; ((\exists a)(a \succsim_D x \;\&\; \neg a \succsim_D y) \vee (\exists b)(y \succsim_D b \;\&\; \neg y \succsim_D b))$.

Taking account of only the first disjunct in the second clause of (c) consider:

(d) $\neg x \succsim_D y \;\&\; (a \succsim_D x \;\&\; \neg a \succsim_D y)$,

From this, together with Axiom 1 we obtain

(e) $y \succsim_D x \vee y \succsim_D a$.

From this $y \succsim_D x$ gives the required result $y \succsim x$ (Def.5.3). So also does $y \succsim_D a$ in conjunction with $a \succsim_D x$ from (d) and Axiom 2.

The second disjunct from the second clause of (c) is dealt with similarly.

Transitivity.

Assume:

$$x \succsim y \ \& \ y \succsim z.$$

This may occur in several ways, as in (a) to (d) below.

(a) $x \succsim_D y \ \& \ y \succsim_D z$,

in which case $x \succsim z$ follows from Axiom 2.

(b) $x \succsim_D y \ \& \ y \succsim_I z$,

i.e. $x \succsim_D y \ \& \ (w)\{(w \succsim_D y \rightarrow w \succsim_D z) \ \& \ (z \succsim_D w \rightarrow y \succsim_D w)\}$.

Putting x as an instance of w gives $x \succsim_D z$ which gives $x \succsim z$ (from Def.5.2).

(c) $x \succsim_I y \ \& \ y \succsim_D z$.

$x \succsim_D z$, and hence $x \succsim z$, follows in a similar way to case (b).

(d) $x \succsim y \ \& \ y \succsim z$,

i.e. $(w)\{(w \succsim_D x \rightarrow w \succsim_D y) \ \& \ (y \succsim_D w \rightarrow x \succsim_D w)\} \ \&$

$(w)\{(w \succsim_D y \rightarrow w \succsim_D z) \ \& \ (z \succsim_D w \rightarrow y \succsim_D w)\}$.

This gives:

$$(w)\{(w \succsim_D x \rightarrow w \succsim_D z) \ \& \ (z \succsim_D w \rightarrow x \succsim_D w)\}$$

from which $x \succsim z$ follows directly.∎

This result shows that, despite the problem of noncomparability, it is possible to order a set of states of physical systems on the basis of an operational test in which some pairs of states are compared directly in the manner characterized in examples such as I.5.1 and I.5.2.

5.8 <u>The Concept *is at least as hot as*</u>

The analysis of the last few sections, leading to the definition of a semi-connected order, reveals that a concept like *is at least as hot as*, if interpreted operationally, is rather more complex than might have been supposed. The added complexity enters the analysis in two places.

One place is the interpretation of the direct relation $\succsim_D$ which now has to refer not only to the nature of the operational tests involved, but also to factors which govern the *possibility* of the test for any pair of states.

The second is in the definition of $\succsim$ given in terms of $\succsim_D$, in which $\succsim$ appears to be endowed with two components of meaning. For example, in the case of temperature the expression $x \succsim y$ may be read as: x is at least as hot as y if either it is at least as hot$_D$ as y or it is at least as hot$_I$ as y. It is the composite term $\succsim$ which corresponds with the ordinary usage of the phrase "is at least as hot as". Both of the locutions "this body is at least as hot as that one" and "this body is at least as hot as it was an hour ago" appear in ordinary discourse. It follows that neither $\succsim_D$ ("is at least as hot$_D$ as") nor $\succsim_I$ ("is at least as hot$_I$ as") means the same as "is at least as hot as". If heat flows from body C to body D when they are in thermal contact, it is a *sign* that C is hotter than D, but it is not the *meaning* of it. Again if when C is in state 1 heat flows from C to D (which is in some fixed state) but when C is in state 2 heat flows from D to C then we have a sign that C in state 1 is hotter than it is in state 2, but not the meaning of it's being so. We might have wondered if the fact that $\succsim_I$ and $\succsim_D$ are related by Def.5.2 gives some sort of warrant for at least according them the same meaning as each other. If this were granted then we could perhaps without much strain extend the same meaning to all three. However a simple numerical example shows that this is not so. Suppose A is a set of numbers and that $x \succsim_D y$ holds if, and only if, $x = (y + 1)$. Then $x \succsim_I y$ holds if, and only if, $x = y$. We would not be inclined to ascribe the same meaning to $\succsim_D$ and $\succsim_I$ in this case. They may be thought of as distinct components of the relation $\succsim$ (as defined in Def.5.3), and if we talk of meanings then we might be prepared to say that the intuitively more complex meaning of $\succsim$ accommodates the meanings of both components, but this in no way justifies conflating the meanings of the components themselves.

It is worth making a point about how this matter relates to the strategy employed in the analysis. Faced with the inadequacy of the original weak order structure of Def.5.1 we have dealt with the problem of noncomparable pairs by making changes in the theory at two levels. One change takes place at (a) the informal level of the interpretation (i.e. in the metalanguage), e.g the change from I.5.3 to I.5.6; the other change is at (b) the formal level of the axioms, e.g the change from the weak order, Def.5.1, to the semi-connected order, Def.5.4. In our treatment so far, the general idea has been to make only the minimum changes at (a), those that proved to be necessary, and to make the bulk of the changes at (b). This has allowed us to keep as close as possible to the spirit of

I.5.3 (and of the other earlier examples). However this is not the only way of dealing with the problem. In principle we can adopt an opposite approach in which we retain the original set of axioms and amend the interpretation accordingly. This means, in effect, adopting the relation $\succsim$ that was *defined* in Def.5.3 as a primitive for the simpler weak order structure of Def.5.1. The difficulty here is that the interpretation is correspondingly complex. The statement of it is rather long and unwieldy, but it goes something like:

I.5.7 The expression $x \succsim y$ holds iff:

either: x and y are states of physically distinct systems and, when the system in state x is in thermal contact with the system in state y, the two are in thermal equilibrium or else heat flows from the system in state x to that in state y,

or: x and y are are states of a single physical system and (i) if when the system in state x is in thermal contact with a second system the two are in thermal equilibrium or else heat flows from the second system into the first, then when the system is in state y and in contact with the second, whose state is unchanged, the two are in thermal equilibrium or else heat flows from the second system into the first, and (ii) if when the system in state y is in thermal with a second system the two are in thermal equilibrium or heat flows from the first system into the second, then when the system is in state x and in contact with the second, whose state is unchanged, the two are in thermal equilibrium or else heat flows from the first system into the second.

Clearly the two solutions are equivalent in that both are based upon the same informal account of the measurement operation. They differ only in how much of that account is made explicit in the formal theory, and how much is left to be spelt out in the specification of the model for the theory. The first leaves the interpretation relatively simple at the cost of complicating the formal theory. The second, by contrast, preserves the simplicity of the theory but requires a more complex interpretation. There is a trade-off from one to the other. The range of choice occurs whenever the fit between a formal structure and an intended model turns out to be incomplete. Some adjustment will be necessary but the decision as to where the adjustment is most appropriately made will depend on what we think formalization achieves. There is a tension between the two approaches described above.

On the one hand a given set of axioms such as that of Def.5.1 may have a strong appeal independently of any peculiar features of an intended application, perhaps because of the simplicity of the axioms or the generality of their application, and for the sake of this it may be worth accepting a more complex interpretation. It may give some motivation for reexamining intuitively held ideas about the intended model on the chance of bringing to light some hitherto unnoticed grounds for adjustment in that quarter. We may come to notice that states of affairs which at first looked quite different can profitably be linked together under some more general description. Such insights are among the fruits of formal analysis.

On the other hand an obvious aim in formalizing an existing theory is to provide a precise expression of the theory as it is intuitively understood. This suggests that simplicity in the interpretation should be a primary goal. It seems desirable that conceptually distinct components of the informal theory should be associated with correspondingly distinct primitive terms in the formal structure. The solution based upon I.5.7 appears to fall down on this score. That the same primitive $\succsim$ should encompass the two relations indicated in I.5.7 is on the face of it contrary to the spirit of this aim.

This does however beg the question of how "conceptually distinct components" are to be recognized. The answer is likely to depend on our metaphysical views on the nature of measurable properties. An operationist is likely to find I.5.6, applied to $\succsim_D$ in Def.5.4, more attractive than I.5.7 for it endows the primitive term $\succsim_D$ with a more clearcut operational significance, and it makes explicit the separation between $\succsim_D$ and the defined ordering relation $\succsim$. For an anti-operationist on the other hand the second scheme based on I.5.7 may have the greater appeal. He will admit that I.5.7 is more complex than I.5.6 but argue that this is a consequence of its being expressed in operational terms. He does not look to I.5.6 for the the meaning of a phrase such as "is at least as hot as" since for him the the phrase refers to a feature of reality that exists independently of the measurement operation. He judges I.5.6 on whether or not it accurately describes an operational test sufficient to pick this feature out. If the test in question is complex and calls for different procedures according to whether states of the same or of different bodies are being compared, so be it. The central point for him is that it is the O-relation $\succsim$ that captures the central underlying concept and that the underlying significance of $x \succsim y$ is the same for any pair x,y, namely an expression of a relation between temperatures (or whatever is the relevant measurable property in a given application).

I do not think that considerations of this sort are compelling for either party. There is a great deal to be said for examining both options, preferably in the order in which they have been described here. The first calls for a deeper analysis of the basis of the measurement process, and requires a more precise expression of it in the formal theory. The second then allows us to make capital of this by revealing how apparently disparate concepts may combine to produce some other familiar notion.

CHAPTER SIX

CONNECTEDNESS, TRANSITIVITY and EXPERIMENTAL ERROR

6.0 Introduction

This chapter, like the previous one, is concerned with problems to do with the failure of the axioms for a weak order. The problem that was considered in Chapter 5, that of noncomparability, has its source in the ontological properties of the physical objects on which the measurement operations are performed. A single object cannot be in two places at the same time, or be in two distinct temperature states at the same time, and so on. We have referred to this type of problem as *essential* failure of the axioms. The type of failure to be considered in this chapter, by contrast, is due not to characteristics of the objects but to those of the measurement operation, and in particular to the intrusion of experimental error. We refer to this variety as *experimental* failure.

We have already given some attention to the subject of experimental error in Chapter 2. It was pointed out that the axioms of Def.1.2 (or Def.5.1) presuppose ideal experimental conditions. In relation to the measurement of weight on a balance, for example, the axioms will be satisfied in all possible tests only if the balance is ideal in construction and behaviour - perfectly symmetrical, perfectly poised, and so on, - and this does not occur in practice. However well constructed a balance may be there is always some departure from an ideal specification and this sets limits to the accuracy and the sensitivity of the tests. In Chapter 2 we were concerned with the consequences of this for the construction of scales. Here we consider the consequences for the formal theory. I wish to discuss how one important type of systematic error leads to failure of the axioms, and to examine possible modifications to the axiom system for dealing with it. In particular I wish to present and discuss the definition of a structure (which I have called a *mutual order*) that appears to meet the problem. It is related to the *semiorder*, a structure first defined by Luce (1956), but it represents a significant development from it and is, I believe, to that extent novel. A major part of the chapter is given to an analysis leading to the definition of a mutual order and to a survey of the ways in which it may be applied. I also include some comment on the fact that there are features of this problem of experimental failure in common with the problem of essential failure.

6.1 A Model for Systematic Error

I shall turn once again to the example of weight measurement, which gives us the following simple model for the type of error I wish to consider.

Suppose we have a balance that is ideal in every respect other than that its arms are unequal in length, in the ratio 1:k, say, as in Fig.6.1 on the following page.

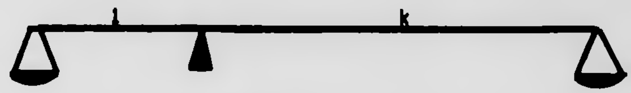


Fig.6.1

k may be less than 1 or greater than 1 but we shall assume that its value is fixed. This produces so-called *systematic* error; but in spite of this it turns out still to be possible to use the defective balance to order a set of objects according to weight, and in this section we investigate how this can be done. Random variation in the value of k would produce *random* error. Since this sort of error is an ineradicable feature of measurement no theory which leaves it out of account can be complete. Nevertheless it is not my intention to pursue this considerably more complex problem here. The features I wish to examine exist beneath the perturbations resulting from random error, and it is clearly more profitable to consider them, at least in the first instance, in an ideal context supposedly free of such perturbations.

The fact that the balance has unequal arms makes it necessary to take account of which object is placed on which pan. Failure to do so quickly leads to unacceptable results. Suppose for example that two objects x and y are nearly equal in weight and, for the sake of this particular point, that k is greater than 1. (It is easier to visualise the outcome of tests if we imagine the departure from ideal conditions to be large - say k = 5, making the right hand arm several times longer than the left.) Now if x is placed in the left hand pan and y in the right then under a standard interpretation of $\succsim$, such as I.5.1, which makes no distinction between the two possible arrangements, the result is:

$$\neg x \succsim y \supset y \succsim x,$$

but with x in the right hand pan and x in the left the result is:

$$x \succsim y \supset \neg y \succsim x,$$

These contradict. To deal with this we must stipulate what arrangement $x \succsim y$ is to refer to. To emphasise the point we change the notation, introducing the predicate L corresponding with which we give the following provisional interpretation of Lxy:

I.6.1 The expression Lxy holds iff, if object x is placed in the *left* hand pan of the balance and object y in the other, the balance remains in equilibrium or else the pan containing x descends.

It is easy to show that ⟨A, L⟩ is, in general, not a weak order. Let us denote the numerical values of weights of objects by n_x, n_y, etc. Then Lxy holds if and only if $n_x \geq kn_y$. Hence we require, for all $x,y,z \in A$:

(a) for connectedness of L: $n_x \geq kn_y$ or $n_y \geq kn_x$,

(b) for transitivity of L: if $n_x \geq kn_y$ and $n_y \geq kn_z$ then $n_x \geq kn_z$.

It is seen immediately that (a) holds in general only if $k < 1$. For (b) we note that the transitivity of $\geq$ (on numbers) guarantees only that:

$$\text{if } n_x \geq kn_y \text{ and } n_y \geq kn_z \text{ then } n_x \geq k^2 n_z.$$

This leads to (b) only if $k \geq 1$. Only for a perfect balance (for which $k = 1$) can both hold together.

In the same way we can introduce a relation R where Rxy has an interpretation obtained by substituting "right" for "left" in I.6.1. Similar arguments show that <A, R> is not a weak order either, the only difference being that the conditions on k for connectedness and transitivity are reversed (assuming that k still refers to the right hand arm as in the above diagram). In fact L and R will have complementary properties; if one is connected but not transitive the other is transitive but not connected. In what follows we shall, without loss of generality, stick to the assumption that $k > 1$, i.e. that the longer arm is on the right in the above diagram, so that L is always the transitive relation and R the connected one of the pair.

There is the question of whether or not L and R are reflexive. Since I.6.1 is a conditional interpretation the same difficulties arise as those described in Section 5.3, and much of the discussion of that section applies again here. In anticipation of numerical models in which Lxy is made to correspond with $x \geq ky$ and Rxy with $kx \geq y$, we are prompted to suppose that L must be irreflexive and R reflexive, but, as before, there is no warrant for any such decision in the results of the operational tests alone. In the next section (6.2) we shall be concerned with the construction of a weak order based on either L or R and it will become clear that for that particular purpose no decision on reflexivity is required. The situation is just the same as that described in the account of the semi-connected order in Section 5.7. We are free to settle the question in whatever way our intuitive promptings suggest without any effect on the results of the subsequent analysis.

We have established that unless k = 1, neither <A, L> nor <A, R> is a weak order, so that neither L nor R is adequate to represent the normally understood relation of "is at least as heavy as". Their deficiencies are simple to describe in informal terms. L is too strong so that Lxy does not hold in some cases where x is heavier than y, those where x is not as much as k times as heavy. R on the other hand is too weak and Rxy holds for cases where y is actually heavier than x, up to k times as heavy. For any pair x, y where neither is as much as k times as heavy as the other we have both of the results:

(i) $\neg Lxy \;\&\; \neg Lyx \quad \text{and} \quad Rxy \;\&\; Ryx.$

That is to say neither L nor R discriminates between objects for which the ratio in weight has a value less than k. Of course this description of the limitations of the system goes well beyond what the operational tests alone entitle us to say. It rests on a particular mathematical theory of beam balances and presupposes that a numerical representation of weight has already been established. At this stage of the analysis we should confine the account to operational concepts, and the most that can be said within this constraint is that there are pairs of objects which are not discriminated by the tests. In this vein it is standard to use one or other of the results in (i) to define a further relation, say:

Def.6.1 For all x,y $\in$ A, Ixy $\equiv$ (Rxy & Ryx).

I is symmetric but, as further tests show, it is not transitive, and therefore it is not an equivalence relation. The standard term for this type of relation is *indifference relation*. Correspondingly a relation of the type of which L is an example is called a *just noticeable difference* relation. (There does not seem to be a recognized term that applies to R alone.) It should again be emphasized that neither of these labels has a numerical connotation and should not be understood to refer to any particular arithmetical relation between objects, which in any case would be subject to change under scale transformations. For instance the geometrical difference of our example would become an arithmetical difference on transformation to a logarithmic scale. (We happen to have chosen an extensive property for illustration but this is incidental at this stage).

Thus, L and R are both insufficient to represent the property of weight. We need to obtain a further relation $\succsim$ that lies somewhere between L and R as it were. We need to find a suitable definition in terms of L, or R, or both.

It is tempting to try simple sentential functions of Lxy and Rxy, but this will not work. It is, for example, simple to show that:

$$x \succsim_{\vee} y \equiv (Lxy \vee Rxy)$$

is connected, but not transitive. Or again, that:

$$x \succsim_{\&} y \equiv (Lxy \,\&\, Rxy)$$

is transitive, but not connected. The relation:

$$x \succsim_{\leftrightarrow} y \equiv (Lxy \leftrightarrow Rxy)$$

is neither transitive nor connected. This third is in any case not very promising from an intuitive point of view since $x \succsim_{c} y$ holds both if x is at least k times as heavy as y and if y is at least k times as heavy as x.

What stands in the way of finding a suitable combination of Lxy and Rxy is that they are not independent. They are trivially related by:

$$Lxy \vee Ryx$$

Intuitively we recognize that the relation $\succsim$ we want for "is at least as heavy as" must be related to L and R by:

$$Lxy \text{ implies } x \succsim y, \text{ and } x \succsim y \text{ implies } Rxy,$$

for any x, y and this gives:

$$Lxy \rightarrow Rxy.$$

This indeed holds for the model of the uneven balance. It will appear as an axiom or a theorem in a satisfactory theory. But the consequence is that Lxy and Rxy cannot be independent sentences in a definition of $x \succsim y$.

6.2 Semiorders

Limits to discrimination in comparison procedures occur universally in measurement and relations with the properties we have ascribed to L and to R are normal. The very fact that measurement is possible despite these limitations, however, indicates that it is possible to establish an order on the basis of just noticeable difference and indifference relations. Goodman described a satisfactory scheme in *The Structure of Appearance* (1951, Chs.IX,X). Subsequently Luce (1956) defined a *semiorder*. This is a structure based on two binary relations, P and I say, where, in the context of his paper, P is a preference relation (which is equivalent to a just noticeable difference relation), and I an indifference relation. Luce gives a set of axioms on the basis of which a weak ordering relation can be obtained from P and I. His definition is (with some notational changes):

Def.6.2 Let A be a set and P and I be two binary relations defined over A. $\langle P, I \rangle$ is a *semiordering* of A [i.e., in our usual format, $\langle A, P, I \rangle$ is a semiorder] if for all w, x, y, z $\in$ A:

1 exactly one of Pxy, Pyx, or Ixy obtains,
2 Ixx,
3 Pxy, Iyz, Pzw imply Pxw,
4 Pxy, Pyz, Iyw imply not both Ixw and Izw.

The import of these axioms is not immediately obvious, but inspection shows that they apply to our system of the uneven balance if Pxy and Ixy are construed in one of several trivially related ways. One is where Pxy corresponds with our (Lxy & ¬Ryx), and Ixy with (Rxy & Ryx), where, in the light of Axiom 2, we must suppose that R is reflexive and, in the light of 1 and 2, that L is irreflexive.

In Def.6.2 P and I are both presented as primitive terms, but Luce does make the point that it is also permissible to define I in terms of P, putting Ixy ≡ ¬(Pxy ∨ Pyx), and we may exploit that here by putting Lxy for Pxy and ¬(Lxy ∨ Lyx) for Ixy. This move does allow the definition to be simplified and so rather than discuss Def.6.2 as it stands it will be convenient to go to a simplified version given by Scott and Suppes (1958). Their definition is:

Def.6.3 A *semiorder* is a relational system ⟨A, P⟩ which satisfies the following axioms for all x, y, z, w ∈ A:

1 Not Pxx,
2 If Pxy and Pzw, then either Pxw or Pzy,
3 If Pxy and Pzx, then either Pwy or Pzw.

Scott and Suppes point out that given a semiorder ⟨A, P⟩ we can derive from it a weak order ⟨A, P_I⟩, where P_I is an indirect relation defined in terms of P of precisely the same type as that employed in the definition of a semi-connected order (Section 5.7). P_I was defined in Def.5.2, (which is renumbered and repeated here for convenience):

Def.6.4 $P_I xy \equiv (w)((Pwx \rightarrow Pwy) \;\&\; (Pyw \rightarrow Pxw))$

In section 5.7 we commented on the operational significance of P_I (or of $\succsim_I$ in the notation of that section) and on how it differs from its direct counterpart P. The main idea was that while a test for Pxy may require direct interaction between the objects x and y, or at least for them both to be present at the same time in an experimental arrangement, a test for $P_I xy$ involves the substitution of one object for the other in an experimental arrangement which is otherwise kept fixed. The difference is that between a test in which objects x and y are placed at the same time on opposite pans of a balance and a test in which x and y are placed successively in one pan while the contents of the other remain unchanged. In the context of section 5.7 the importance of indirect tests was that they avoided problems of connectedness arising from logical difficulties in setting up direct tests. But as we shall now see this same notion is exploited in the application of the semiorder to experimental error.

The following is a theorem.

Th.6.1 If the structure $\langle A, P \rangle$ is a semiorder and P_I is a binary relation on A defined from P as in Def.6.4 then the structure $\langle A, P_I \rangle$ is a weak order.

The proof of this is fairly elementary and the authors omit it from their paper, but it is important for some of the following analysis to examine the precise role of each of the axioms, and to this end we spell it out on the following page.

Proof: Connectedness.

Assume:

(a) for some x,y: $\neg P_z xy$ & $\neg P_z yx$,

i.e. $\neg(w)((Pwx \rightarrow Pwy)$ & $(Pyw \rightarrow Pxw))$ & $\neg(w)((Pwy \rightarrow Pwx)$ & $(Pxw \rightarrow Pyw))$.

The first conjunct of this gives:

(b) $(\exists a)((Pax$ & $\neg Pay) \vee (Pya$ & $\neg Pxa))$,

and the second gives:

(c) $(\exists b)((Pby$ & $\neg Pbx) \vee (Pxb$ & $\neg Pyb))$.

The proof now falls into two parts.

The first consists in showing that the conjunction of the first disjuncts from (b) and (c) and, similarly, the conjunction of the second disjuncts violate Axiom 2. The first pair for example give:

$(Pax$ & $\neg Pay)$ & $(Pby$ & $\neg Pbx)$, i.e. Pax & Pby & $\neg(Pay \vee Pbx)$

which contradicts Axiom 2.

The second part of the proof consists in showing that the conjunction of the first disjunct of either (b) or (c) with the second disjunct of the other violates Axiom 3. For example the first disjunct from (b) and the second from (c) gives:

$(Pax$ & $\neg Pay)$ & $(Pxb$ & $\neg Pyb)$, i.e. Pax & Pxb & $\neg(Pay \vee Pyb)$,

which contradicts Axiom 3.

Thus assumption (a) leads to a contradiction.

Transitivity.

Follows immediately from Def.6.4 alone.∎

Now inspection shows that Axs.2 and 3 of Def.6.3 are satisfied by both of our relations L and R. However given that it is usual to count L but not R as irreflexive, the presence of Ax.1 indicates that the system is intended to apply only to L, i.e. only to a just noticeable difference relation, so that Th.6.1 applies to the structure $\langle A, L_I \rangle$ but not to the corresponding structure $\langle A, R_I \rangle$. Now in the light of comments we have made earlier about the irrelevance of the question of reflexivity it may be wondered if this restriction is necessary. We note that the foregoing proof makes no appeal to Axiom 1; in so far as we are concerned merely with establishing the existence of a weak order this axiom is redundant. It is quickly shown to be independent of the other two. Consider the numerical relational structures $\langle Re, > \rangle$ and $\langle Re, \geq \rangle$. Axs.2 and 3 hold for both, whereas Ax.1 holds only for the former. If we discard Ax.1 Th.6.1 applies to $\langle A, R_I \rangle$ as well.

The question now arises of what these weak orders $\langle A, L_I \rangle$ and $\langle A, R_I \rangle$ are. Are they the same as the weight order? Is L_I (or R_I) an adequate represenattion of the relation "is at least as heavy as"?

The most immediate answer is that L_I (or R_I) is an *approximate* version of the ideal, direct, weak order, relation that we suppose to be obtainable with a perfect balance, one for which k is exactly 1. Suppose that we represent this ideal direct relation by W. (In the limit of k=1 both L and R would reduce to W.) Now if W is a weak order then W and W_I are equivalent - the result $Wxy \leftrightarrow W_Ixy$ follows very simply from the axioms for a weak order. But $Wxy \leftrightarrow L_I xy$ does not hold in general. We recognize intuitively that:

$$Wxy \text{ implies } L_I xy$$

must hold, but the reverse:

$$L_I xy \text{ implies } Wxy$$

is not guaranteed. For suppose that for a particular pair, a and b, we have ¬Wab. Then we hope for $\neg L_I ab$ but for this there must be some other element, c say, such that either (Lca & ¬Lcb) or (Lbc & ¬Lac) holds. Whether or not this is so is contingent upon the membership of the set A. If A contains no such element the procedures will fail to discriminate between a and b. In this event the experimental order based upon L_I will be cruder than the ideal order of W that we are after. The failure of L_I to discriminate in certain cases will lead to the deletion of gaps that would appear in the ideal order. It is in this sense that the experimental order is an approximation to the ideal. It is clear that the more finely distributed the set is, the greater the chances of discriminating between pairs of elements, and so reducing this type of error. In particular, if there were available a continuum of objects, with weights distributed over a continuum of values, it would be possible in principle, so we suppose, to use L_I to reproduce exactly the order W would give. On this view, $\langle A, L_I \rangle$ approaches $\langle A, W \rangle$ as a limit as the membership of A is increased indefinitely and L_I approaches W.

This argument presupposes the theory of an ideal beam balance and assumes that if a perfect instrument were constructed, with k exactly equal to 1, it would behave as our informal arguments have supposed. Maybe it would, but it is not logically necessary that it should. It is possible that, even with a symmetrical balance, W is an indifference relation. It could be the case for example that moments of forces obeyed some kind of quantization principle so that two objects must differ in weight by more than a certain threshold value before the equilibrium of the balance is disturbed. Indeed it could be that if experimental error is reduced below some as yet unknown level, precise measurement will reveal effects of this kind. Given that we cannot completely eliminate experimental error we can never be in a position to be sure that this is not actually the case. We cannot be sure that the supposed underlying ideal weak ordering relation does exist. And if it does not, then even with a perfect balance at our disposal, we would still need to appeal to an indirect relation W_I to obtain a weak order and W_I not W would be the underlying relation to which L_I approximates. This is not to claim that W_I and L_I are incorrigible. The point I wish to establish is simply this. Ordering a set by a relation representing a direct mutual interaction imposes more stringent conditions on the test operation than ordering by the alternative indirect method. This is some reason for regarding the indirect relation as the more fundamental from the point of view of establishing the existence of a measurable property.

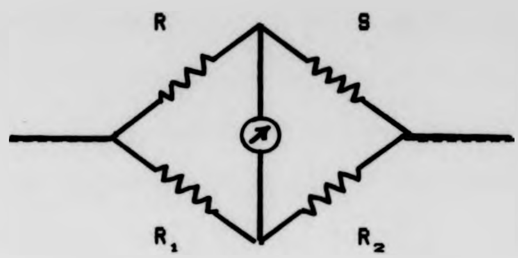
Thus to summarise, the semiorder copes with the problem of the uneven balance, (and with kindred problems in other systems where the same sort of experimental error occurs), by replacing a binary relation P, rendered unsuitable by shortcomings in the measuring operation, by its indirect derivative P_I. Any remaining error is then attributable to deficiencies in the distribution of the set of objects along the order. It is possible in principle to reduce these errors to an indefinite extent by suitably augmenting the set.

6.3 Method of substitution

The technique by which the semiorder deals with the problem of experimental error is more than a mathematical contrivance. It corresponds more closely than might at first be apparent with operations employed in practice. There are very many cases of routine procedures in which greater precision is obtained by comparing two objects by an indirect method rather than by a direct method. We describe a particular example.

Suppose we wish to adjust a variable resistor R until the value r of its resistance is equal to that of some resistance standard r_s, say. Typically we make use of a Wheatstone Bridge arrangement. (An account of the theory of the Wheatstone Bridge, together with some comment on the substitution method can be found in Edwards (1971, p.10).) A *direct* method would use the cicuit depicted in Fig.6.2 on the following page.

Fig.6.2



R is the variable resistor and S the standard. The circuit contains two further resistors whose values r_1 and r_2 are nominally equal, giving a a nominal value of 1 for the ratio $k = (r_1/r_2)$. The variable resistor R is adjusted until the bridge is balanced, that is until no current flow is registered by the meter at A. Standard circuit theory gives for the balance condition:

$$r = (r_1/r_2) \times r_s = k \times r_s$$

As this equation shows, the accuracy of the setting of the value r to r_s is dependent upon how close to 1 the value of k in fact is. Any departure of the value from 1 will be revealed if the two resistors R and S are interchanged when in general it will be found that the bridge is no longer in balance. In this respect the behaviour of the circuit as a 'resistance balance' is entirely similar to that of the weight balance, with arms of lengths in the ratio 1:k, that was described in an earlier section. The accuracy of the method is limited by the imperfection of the bridge. A standard method of dealing with this problem is to employ the arrangement of the following Fig.6.3.

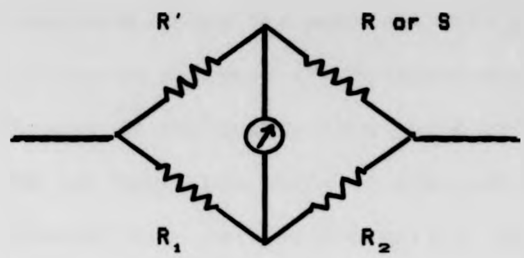


Fig.6.3

There R' is another variable resistor that is put in place of R. First the bridge is balanced by adjusting R'. Then, without any further change to R', R is substituted for S and adjusted until the bridge is again balanced. R (now fixed at this value) and S can then be interchanged at will without any change in the condition of the bridge. We may then conclude that $r = r_s$, and this result is independent of the values of any other components in the circuit; it holds for any value of k.

The principle of this method is in line with the analysis of the semiorder. We may obviously describe it as an indirect comparison of R and S by means of a direct comparison of each of them with R'. Note that the method owes its success to the fact that R' is a variable resistor. This provides in effect a notionally continuous set of resistors out of which we select the one required for the most precise comparison of R and S.

But there is a more general way of looking at the matter. As the analysis shows the value of the resistance of R', or, to talk about it in pre-calibration terms, the position of R' relative to R and S in the overall order, is irrelevant. From the point of view of this technique it need not even be regarded as a member of the ordered set. The function of R' is simply to form part of the system into which the resistors under test are inserted. We can count the whole of the circuit of diagram (b) - except for R and S themselves - that is not only R' but the other resistors, detector and the rest, as a complete system, or *environment*. On this view we compare R and S by *comparing the outcomes when one is substituted for the other in an appropriate, fixed environment.* I now wish to argue that we may apply this alternative view generally across the whole field of measurement and that there are substantial advantages to be gained from doing so. I shall begin by describing how an alternative formal system which exploits this viewpoint can be developed from the semiorder structure. The essence of this alternative system is that it separates the objects to be ordered from the environments which provide the contexts in which the behaviour of the objects is compared.

6.4 Definition of a Demisemiorder

The development can best be explained in the following way. We return to Def.6.4 and rewrite $P_I xy$, making use of an elementary theorem of predicate calculus, as:

$$P_I xy \equiv (w)(Pyw \rightarrow Pxw) \mathbin{\&} (w)(Pwx \rightarrow Pwy).$$

The clauses of the conjunction can be separated in an obvious way, and we now define two relations P_J and P_K, components of P_I, as follows:

Def.6.5 (i) $P_J xy \equiv (w)(Pyw \rightarrow Pxw)$, and

(ii) $P_K xy \equiv (w)(Pwx \rightarrow Pwy)$,

so that:

$$P_I xy \equiv (P_J xy \ \& \ P_K xy).$$

Both P_J and P_K are reflexive and transitive, again as a matter of definition. Furthermore, as examination of the proof given in the last section readily shows, each of them is connected so that both $\langle A,P_J\rangle$, and $\langle A,P_K\rangle$, are weak orders. The interesting point is that this can be established on the basis of Axiom 2 of Def.6.3 alone. This single axiom is sufficient. This comes about because of the difference in operational significance of P_I on the one hand and P_J and P_K on the other. $P_I xy$ refers both to cases where object x is inserted in one test position of the measurement system (e.g. the left hand pan of the balance or the left hand branch of the Wheatstone Bridge) and to cases where it is inserted in the other position, and the same for y. P_J and P_K however refer only to cases for which any given element is restricted to one position or the other. Axiom 3 in the Def.6.3 governs the behaviour of the system in tests involving the transfer of an object from one side to the other, and so it is not required if no such tests are referred to. Axiom 2 is independent of any such interchange.

The function of Axiom 3 is, as it were, to couple the two relations P_J and P_K together. It is this axiom that ensures that the two orderings of A that they produce are compatible so that P_I, their combination, also gives a weak order. Axiom 3 incorporates information about a more complex aspect of the behaviour of the measurement system than does Axiom 2. But, as the above arguments show, we can dispense with this information and still obtain an order. As a first step, therefore, we define a weaker structure than a semiorder, which we shall call a *demisemiorder*, based on a single axiom:

Def.6.6 Let A be a set and P a binary relation on A. The relational structure $\langle A, P\rangle$ is a *demisemiorder* iff, for all u, v, x, y $\in$ A, the following axiom holds:

$$(Pxu \;\&\; Pyv) \rightarrow (Pxv \vee Pyu).$$

We may now give a formal summary of the independent results for P_J and P_K in the following theorem.

Th.6.2 If $\langle A, P\rangle$ is a demisemiorder then the relational structures $\langle A, P_J\rangle$ and $\langle A, P_K\rangle$, where P_J and P_K are relations defined on A as in Def.6.5 are weak orders.

This weaker structure generates two orders out of tests conducted on an uneven balance, (or other similarly defective equipment). Each is an approximation to an ideal, where once again the degree of approximation is governed by the distribution throughout the set of objects. What has been relinquished in the move from Def.6.3 to Def.6.6 is the guarantee that P_J and P_K give the same order, or give approximations to the same ideal order. This structure leaves that question open.

The link with the problem of essential failure of connectedness is now easy to state. Suppose we have a situation where the set A is divided into two subsets such that all the elements of one subset are restricted to appearing on one side of the system only and the elements of the other subset similarly restricted to the other side. That is, no two members of the same subset can be compared directly. Then Axiom 3 of Def.6.3 is not invoked; only half of P_I would be needed for any given element, depending on which subset it belonged to, and in this situation Def.6.3 reduces to Def.6.6. By way of illustration we revert to our example of the uneven balance and consider the following rather fanciful situation.

Suppose that for some reason it is impossible to transfer objects from one arm of the balance to the other. It could be extraordinarily long, or there could be an impassable barrier through which the beam of the balance passes but which otherwise separates left from right. There is an operator at each end with his own set of objects that he wishes to order by weight. They cooperate in informing each other which particular member of their set is on the balance at a given time as they work through an exhaustive series of tests. Each is then able to decide for any pair of his own set that one is at least as heavy as the other on the basis that it tips the balance whenever the other does. What the tests cannot do is determine how any members of one set compare for weight with those of the other. For that we require extra information or assumptions, the value of k together with the theory of beam balances, for example. In the absence of such information the best that can be done is to come to a conventional decision if one is required. This procedure can be formalized in a scheme to be described in the next section.

6.5 Definition of a mutual order

We begin by defining a structure weaker than a semiorder.

Def.6.7 Let A and B be sets and P a binary relation defined on $A \times B$. The relational structure $\langle A \cup B, P \rangle$ is a *mutual order* iff, for all $x, y \in A$ and $u, v \in B$, the following axiom holds:

$$(Pxu \;\&\; Pyv) \rightarrow (Pxv \vee Pyu).$$

The idea of the term *mutual order* is that on the basis of this single axiom, it is possible to order both sets at once. We can express this formally in a theorem:

Th.6.3 If $\langle A, B, P \rangle$ is a mutual order and two new relations, P_J on A and P_K on B are defined by:

(i) $P_J xy \equiv (w)(Pyw \rightarrow Pxw)$,

(ii) $P_K uv \equiv (z)(Pzu \rightarrow Pzv)$,

then the relational structures $\langle A, P_J \rangle$ and $\langle B, P_K \rangle$, are weak orders.

As far as the present example is concerned A and B stand for the sets of objects at each end of the balance. Either L or R may be put in place of P. Notice that if the barrier is removed so that objects can be transferred freely, we obtain A = B and the mutual order reduces to the demisemiorder of Def.6.6. This example is somewhat contrived; it really comes into the category of experimental failure, as described at the beginning of this chapter. We can give more serious examples, involving essential failure, of the application of this structure.

6.6 Examples of mutual orders

(a) Suppose we have a set of solid bodies and a set of liquids which are to be ordered by density. P is interpreted so that Pxu means "if x is placed in u it sinks". P is not a weak ordering relation since for reasons deriving from the nature of the objects involved it is not connected on the total set. It is not possible for one solid to sink in another solid. Given two solids x,y we can have neither Pxy nor Pyx. And for the sake of illustration let us suppose that all the liquids are pairwise miscible, which is counted to give ¬Puv for any pair u,v. Thus all the solids belong to A and all the liquids belong to B. Now the axiom of 6.6 requires no more than that if solid x sinks in liquid u and solid y sinks in liquid v then either x sinks in v or y sinks in u. That this is in fact the case means that we can order such a set of solids, and at the same time order the set of liquids, according to what we recognize intuitively as density, on the basis of the relation P. One solid is counted at least as dense as a second if it sinks in any liquid in which the second sinks. One liquid is counted at least as dense as a second if any solid that sinks in it sinks in the second.

What this procedure does not do is determine how the densities of solids are related to the densities of liquids. It does not even give any warrant for claiming that the two sets have been ordered according to the same property, that the meaning of "density" as applied to liquids is the same as when it is applied to solids. To come to a judgement on this, to relate the sets A and B, we again need further information. Intuitively we recognize that if a solid sinks in a liquid then it is denser than the liquid, and we are inclined to interpret the term Pxu accordingly. But this is based upon theory about weight, volume, flotation and so on. In the absence of such information there is no way of telling whether or not P, P_J and P_K all signify the same measurable property. The possibility that all three have different meanings is seen in a simple model. Let A be the set of parents and B the set of children in a (strictly monogamous) family, and P the relation "is a parent of". Then P_J stands for "is identical with or is the spouse of", while P_K stands for "is identical with or is a sibling of". (The weak orders here degenerate to equivalence classes.)

Notice that the situation changes if some pairs of liquids are immiscible so that one may sink in a layer to the bottom of the other, or if we allow tests in which a sachet filled with one liquid is immersed in another. Then at least some of the liquids are members of A as well as of B. Whether or not P, P_J and P_K signify the same relation is now testable. This possibility arises whenever $(A \cap B)$ is non empty. These two cases may be distinguished by referring to the first - where A and B are disjoint - as a *strict* mutual order.

To get some feel for how the axiom has worked in this example consider what obtains if we interpret Pxu instead to mean "if body x is placed in liquid u it dissolves". In general the axiom will not hold for this interpretation and so we do not get an order of solubility.

(b) The previous example indicates that it is not necessary for the members of A to be the same kind of entity as those of B. Suppose that instead of a chemical balance we use a steelyard to order a set of weights of the sort shown in this Fig.6.4. The pan on the left is fixed but the single weight on the right can be moved to any of a set of positions $k_1, k_2, \ldots\ldots$ on the right.

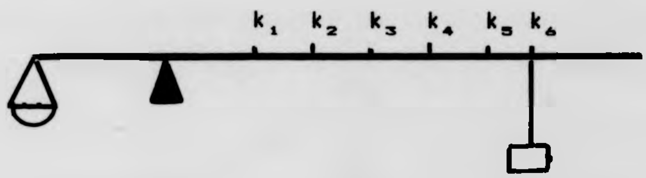


Fig.6.4

Then let A be a set of bodies, and let B be the set of positions. Then P_J orders the bodies by weight. P_K orders the positions according to distance from the fulcrum.

(c) Suppose A is a set of pendulums and B a set of time periods, i.e. a set of intervals on a time axis, and the interpretation of Pxu is: "if pendulum x starts a swing at the start of period u, the end of the period occurs at or before the pendulum completes a cycle". This application presupposes some means of individuating time periods. This can be done by reference to observable events. Among these may be counted any particular swing of any pendulum in the set A. The time interval defined by that swing may therefore be a member of B. But this is not to say, of course, that A and B have members in common. The pendulum which is a member of A is a different entity from one of its swings which is a member of B. This type of application is important for its bearing on the problem of calibrating a metric.

(d) Suppose that A is a set of resistors and B a set of Wheatstone Bridge arrangements of the kind portrayed above in section 6.3, with one arm unoccupied as in Fig.6.5.

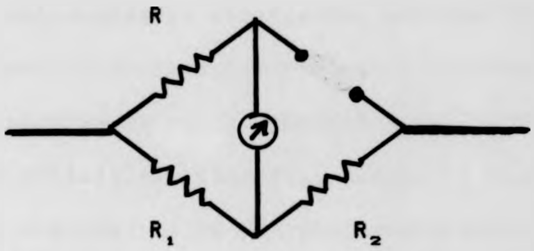


Fig.6.5

The detector is an ammeter which need indicate no more than when a current is flowing through it and in which direction, say a centre-zero instrument whose needle moves either to left or right, but without any calibration. A resistor from set A can be connected into the bridge between a and b. Different bridges are obtained by altering the components in the other arms in any way whatever. Pxu is to mean: "when resistor x is connected into arrangement u the needle on the meter moves to the left". This gives an ordering of the set of resistors by the relation P_3.

The set B is also ordered. That is, the set of bridge arrangements is ordered, by P_K. In this example it is natural to view each circuit arrangement as an environment into which the objects of set A may be inserted for observation. This viewpoint could be adopted in any of the applications of a strict mutual order. In example (a) we may perhaps think of the liquids as environments in which the behaviour of the solids can be observed. In (b) the different settings of the steelyard, or for that matter settings of any number of steelyards, provide environments. For (c) we can think of a particular time interval as a temporal environment in which a pendulum acts, and so on. In general A is the set of objects to be ordered and B a set of test environments, though in many cases the distinction between the two may be somewhat arbitrary. A natural reading of Pxu is then: "object x in environment u gives rise to the outcome P". The axiom in Def.6.7 may then be read "If outcome P is obtained with object x in environment u, and with object y in environment v, then it is also obtained in at least one of the cases x in v and y in u".

CHAPTER SEVEN

CONCATENATION and EXTENSIVE MEASUREMENT

7.0 Introduction

In Chapter 5 we considered certain problems in the application of the standard weak order structure $\langle A, \succsim\rangle$ (Def.1.2) to the measurement of physical attributes. The problems were due to fundamental differences between physical objects and numbers. We needed to take account of the fact that the nature of physical objects imposes constraints on the empirical relations that can exist among them, constraints that do not arise with regard to the arithmetical relations among numbers to which, as the basis for measurement, the empirical relations among physical objects are taken to correspond. As we saw, an important consequence of the existence of these constraints is that, in general, the ordering relation underlying a measurable property cannot be identified with a relation expressing the outcome of a direct operational test comparing two objects.

In this chapter we turn to extensive measurement, and consider problems in a similar vein that arise in that area of the formal theory. In particular we will examine difficulties in the structure $\langle A, \succsim, \circ\rangle$ of Def.1.3, which was cited as a typical example of a structure of extensive measurement. The presence of a concatenation operation in a structure such as this increases the scope and complexity of problems of the sort already encountered with Def.1.2, and given the central position of structures like $\langle A, \succsim, o\rangle$ in the theory of measurement, it is important to extend the investigation of Chapter 5 in this direction.

In Sections 7.1 to 7.3 we describe the problems in question and propose some possible ways of dealing with them within the framework of a standard definition of $\langle A, \succsim, o\rangle$. Definitions of a *partly connected order* (Def.7.3) and a *proper concatenation structure* (Def.7.5) are given. Section 7.4 contains a detailed discussion of concatenation procedures centred on the question of associativity and commutativity of the concatenation operation. Section 7.5 shows an extensive measurement structure formulated in Boolean terms. Section 7.6 contains an account of a further problem due to constraints on the concatenation operation and a possible solution is suggested in Section 7.7.

7.1 Axioms for extensive measurement

We start by quoting again the following definition, from Krantz *et al* (1971), which was first introduced in Chapter 1 as Def.1.2. Though there are a number of variants on this definition to be found in the literature, it is reasonable to view it as a standard definition of a structure of extensive measurement and to take it as a basis for our investigation. It appears on the following page and is renumbered as Def.7.1.

Def.7.1 Let A be a nonempty set, $\succsim$ a binary relation on A, and $\circ$ a closed binary operation on A. The triple $\langle A, \succsim, \circ\rangle$ is a *closed positive extensive structure* iff the following axioms are satisfied for all $x, y, z \in A$:

1 $\succsim$ is a reflexive, transitive and connected relation.
2 $x \circ (y \circ z) \sim (x \circ y) \circ z$.
3a $x \succsim y$ iff $x \circ z \succsim y \circ z$.
3b $x \succsim y$ iff $z \circ x \succsim z \circ y$.
4 $x \circ y \succ x$.
5 If $x \succ y$, then for any $v, w \in A$, there exists a positive integer n such that $nx \circ v \succsim ny \circ w$, where nx is defined inductively as: $1x = x$, $(n + 1)x = nx \circ x$. [i.e. nx results from the concatenation of n replicas of x.]

We made some explanatory comments about the significance of these axioms in Chapter 1. The first is of course the condition for $\succsim$ to give a weak ordering of the set A, and it was problems associated with this axiom that provided the subject matter of Chapter 5. A large part of that discussion applies again here, though because of a complication arising from the concatenation operation the solutions proposed in that chapter will be found to be insufficient to deal with the corresponding problems in this structure.

The other axioms involve the concatenation operation explicitly. I shall sketch the reasons why the inclusion of this operation in the structure leads to added problems.

According to the closure condition on A, for any two elements x, y of A, there is a third element $x \circ y$, which it is usual to interpret as the object obtained by combining, in some specified way, the objects that are represented by x and y. It is a third object of which those represented by x and y are constituents. For example if x and y are bodies that can be placed separately on one pan of a balance, x o y is the pair of them placed on the pan together. If x and y are rigid rods, x o y is the composite rod formed by abutting the two end to end. In general x o y is assumed to represent a further physical object having the same measurable property that is attributed to its components x and y. Again the composite object x o y can be combined with another object z to give a further object (x o y) o z and so on. In addition, we must take account of the fact that the closure of o yields, say, y o x as well as x o y, and x o (y o z) as well as (x o y) o z. It is usual to assume that pairs of terms such as these represent objects constructed from the same constituents but assembled in a different order. [Whether or not x o y and y o x, or x o (y o z) and (x o y) o z, represent the *same* body, that is whether or not o is commutative and associative is a question that is left open by Def.7.1. It is a question that proves to be of some significance and we will have reason to pay close attention to it in due course, but the following remarks apply irrespective of the answer.]

It follows from all this that there will be among the members of A pairs such as x o y and x, x o y and x o z, x o y and y o x, and so on, that is, pairs of objects either where one object is part of the other, or where both have some part in common, or where both are composed of the same constituents. The objects x and y need not be physically joined to form the object x o y, nor even be in contact; it may be sufficient for them to be present at the same time, e.g. in the same pan of a balance.

But even so, two objects represented by x o y and x, or by x o y and x o z, and so on overlap in physical content. They are pairs of objects for which the existence of one is not independent of the existence of the other. Concatenation adds richly to the stock of those pairs of members of A of which neither can be manipulated independently of the other, and their existence is troublesome in two ways.

First, they are new examples of what were described in Chapter 5 as noncomparable pairs. No such pair is connectable under a simple direct operational interpretation of $\succsim$ and so further problems of the sort already dealt with at some length in Chapter 5 will arise. According to Axiom 1, for example, the following hold for any $x,y,z \in A$:

$$(x \circ y \succsim x) \vee (x \succsim x \circ y),$$
$$(x \circ y \succsim x \circ z) \vee (x \circ z \succsim x \circ y)$$
$$(x \circ y \succsim y \circ x) \vee (y \circ x \succsim x \circ y)$$

and examples could be multiplied. They show that the relation $\succsim$ is to hold between two objects one of which is part of the other, or which share a common constituent, or which are made up from the same constituents. But if $x \succsim y$ has a simple, direct operational interpretation - one requiring the objects x and y to be simultaneously present, such as being placed on opposite pans of a balance at the same time - then expressions such as $(x \circ y \succsim x)$ and $(x \circ y \succsim x \circ z)$ refer to states of affairs that are not physically realizable and so cannot be tested. Inspection shows that problematic expressions like these occur in one guise or another in applications of all the other axioms. As before we may pose the problem of noncomparable pairs in the following general terms. Given that the truth values of expressions of this sort cannot be determined by the outcome of direct operational tests, according to what alternative principles are they to be determined?

The second problem is to do with the meaning of the term $x \circ y$ when x and y stand for bodies that have a common constituent. How are we to construe terms like $(x \circ y) \circ (x \circ z)$, $(x \circ y) \circ x$, or $x \circ x$? The closure of $\circ$ requires that there be members of A to which such composite terms refer, but what sense is to be made of them? A physically correlated pair cannot be concatenated, at least not in the straightforward sense that applies, as in the illustrative examples we gave above, to wholly physically distinct objects. What is to count as the concatenation of two physical objects that have some part in common, or as the concatenation of an object with itself? This is another point at which the properties of physical objects diverge from those of numbers. Numbers may be added whether they are distinct or not. In the world of numbers "1 + 1" makes as good sense as "1 + 2"; in the world of physical objects we can add one brass weight to a second distinct brass weight but we cannot add a brass weight to the same brass weight. We can refer to this as the problem of constructibility. Its consequence is that under the usual sort of interpretation of $\circ$ there are terms in the formal language that have no reference. Again we face the question of how to determine truth values of sentences in the language that contain terms that have no direct operational significance.

In the rest of this chapter we consider the ramifications of these twin problems of comparability and constructibility.

7.2 Concatenation and Comparability

We recall from Chapter 5 the general approach that was adopted for problems of this kind. It was to seek to define, in terms of the original direct relation $\succsim_D$, an ordering relation $\succsim$ that applies both to comparable and to noncomparable pairs, and to find a reasonably simple set of axioms sufficient to ensure that $\succsim$ is a weak ordering relation. The same approach is followed successfully here, though the success is qualified by the need to impose stronger associativity and commutativity conditions on the concatenation operation than those that are usually assumed. In what follows we draw as far as possible on parallels with the ideas of Section 5.6, with the aim of giving some continuity with the earlier material, as well as identifying the important additional features of the problem that are peculiar to a concatenation operation.

7.2.1 Definition of an Ordering Relation

We consider again the example of an ordering relation for weight. As with the cases dealt with in Chapter 5, we begin with a suitable categorical interpretation for the D-relation:

I.7.1 The expression $x \succsim_D y$ holds iff it is possible to place objects x and y on opposite pans of a balance at the same time and, if objects x and y are placed on opposite pans of a balance, the balance remains in equilibrium or else the pan containing the object x descends.

The next step is to formulate a suitable definition of the O-relation, $\succsim$, which must satisfy the following conditions. Firstly, it must apply to cases covered by $\succsim_D$, that is we must have:

(a) $x \succsim_D y \rightarrow x \succsim y$.

Secondly, as we intuitively recognize, $x \succsim y$ is to hold in the following cases:

(b) x and y have exactly the same constituents, or
(c) x contains y as a constituent, or
(d) x and y have one or more, but not all, constituents in common, and if the object that remains after removing the common constituents from x is placed in one pan of a balance, and the object that remains after removing the same constituents from y is placed in the other, the balance remains in equilibrium or else the pan containing the residue of x descends.

Clauses (b) to (d) cover (exhaustively) the noncomparable pairs that are missed by I.7.1. and therefore missed by (a). The list (a) to (d) is not intended to be an informal version of the sought after definition of $\succsim$. It is a set of criteria by which to assess the adequacy of any proposed definition. We wish to define $\succsim$ in terms of operational concepts, that is in terms of $\succsim_D$, alone but these rules are expressed in terms of a mixture of operational concepts and others relating to the structure of individual objects. The complexity is reflected in the variety of criteria involved. The test for whether or not $x \succsim y$ holds varies according to which of the categories (a) to (d) apply. (a) and (d) involve operations with the balance whereas for (b) and (c) the truth value of $x \succsim y$ is determined simply by inspection of the structure of the objects x and y.

We seek a definition of $\succsim$ in terms of $\succsim_D$ that captures these, and only these, cases.

It is worth pointing out first of all that the definition that was successful in dealing with noncomparability between states of the same body, as in the examples in Chapter 5, is not adequate here. We recall

Def.5.3

$$x \succsim y \equiv x \succsim_D y \vee (w)\{(w \succsim_D x \rightarrow w \succsim_D y) \,\&\, (y \succsim_D w \rightarrow x \succsim_D w)\}.$$

Now inspection shows that this works for the cases covered by clauses (a), (b) and (c), but breaks down for those cases described by (d). To see this suppose, for example, that $x = a \circ b$ and $y = a \circ c$ and also that $b \succsim_D c$ hold. We then require $x \succsim y$ to hold. But suppose further that for some $w = c \circ d$ it also holds that $w \succsim_D x$. This w has a constituent in common with y and so $w \succsim_D y$ cannot hold. Hence the definition fails.

A relation that does appear to meet all the cases (a) to (d) is:

Def.7.2 for all $x, y, w \in A$

$$x \succsim y \equiv (w)\{(w \succsim_D x \,\&\, y \succsim_D w) \rightarrow (x \succsim_D w \,\&\, w \succsim_D y)\}$$

This may not be intuitively obvious but it may perhaps be rendered more transparent if the *definiens* is recast in the following way:

$(w)\{[(w \succsim_D x \,\&\, y \succsim_D w) \rightarrow x \succsim_D w] \,\&\, [(w \succsim_D x \,\&\, y \succsim_D w) \rightarrow w \succsim_D y]$

which gives:

$(w)\{[(y \succsim_D w) \rightarrow (x \succsim_D w \vee \neg w \succsim_D x] \,\&\, [(w \succsim_D x) \rightarrow (w \succsim_D y \vee \neg y \succsim_D w)]\}$

This may be paraphrased: anything that fails to tip the balance against object y also fails to tip it against object x, or else is not comparable with x at all; and anything that tips the balance against x also tips it against y, or else is not comparable with y at all. Very roughly, any w that is comparable with both x and y, and so can be used to compare x and y by an indirect means, shows the appropriate behaviour. It is perhaps slightly disconcerting that the definition, as given in the first line, is a conditional that, in cases where it holds, does so by virtue of the antecedent's always being false, except only for cases where x and y are equivalent for then the antecedent will be true for any (comparable) w that is itself equivalent to each of them. Nevertheless, technically at least, Def.7.2 is perfectly adequate.

A relation that does appear to meet all the cases (a) to (d) is:

Def.7.2 for all $x,y,w \in A$

$$x \succsim y \equiv (w)((w \succsim_D x \;\&\; y \succsim_D w) \rightarrow (x \succsim_D w \;\&\; w \succsim_D y))$$

This may not be intuitively obvious but it may perhaps be rendered more transparent if the *definiens* is recast in the following way:

$(w)([(w \succsim_D x \;\&\; y \succsim_D w) \rightarrow x \succsim_D w] \;\&\; [(w \succsim_D x \;\&\; y \succsim_D w) \rightarrow w \succsim_D y]$

which gives:

$(w)([(y \succsim_D w) \rightarrow (x \succsim_D w \vee \neg w \succsim_D x] \;\&\; [(w \succsim_D x) \rightarrow (w \succsim_D y \vee \neg y \succsim_D w)])$

This may be paraphrased: anything that fails to tip the balance against object y also fails to tip it against object x, or else is not comparable with x at all; and anything that tips the balance against x also tips it against y, or else is not comparable with y at all. Very roughly, any w that is comparable with both x and y, and so can be used to compare x and y by an indirect means, shows the appropriate behaviour. It is perhaps slightly disconcerting that the definition, as given in the first line, is a conditional that, in cases where it holds, does so by virtue of the antecedent's always being false, except only for cases where x and y are equivalent for then the antecedent will be true for any (comparable) w that is itself equivalent to each of them. Nevertheless, technically at least, Def.7.2 is perfectly adequate.

7.2.2 Definition of a Partly Connected Order

The final stage of the problem is to state axioms sufficient for $\succsim$ of Def.7.2 to give a weak ordering of A. It was mentioned in Chapter 5 that there is always a trivial solution to this. It is simply to state formally that $\succsim$ satisfies this condition. However, as always we want the axioms to be as simple and intuitively clear as possible, and the bald statement that the relation $\succsim$ of Def.7.2 is connected and transitive hardly meets this requirement and the aim is to produce something simpler than that.

Perhaps it should first be pointed out that, for the same reason that the relation $\succsim$ of Def.5.3 is inadequate for the present problem, the set of axioms that were formulated for that particular relation, namely those defining a semi-connected order, Def.5.4, will not do either. Ax.1 of that structure requires that the relation of noncomparability is transitive. The relations indicated in (b) and (c), respectively that of one thing's being composed of the same constituents as another, and that of one thing's wholly containing another, do indeed satisfy this, but the relation indicated in (d) again proves to be a problem. This relation, of one thing's having one or more constituents in common with another, is not transitive. We therefore need something different from Def.5.4.

The following definition of what we may term a *partly connected order* appears to be sufficient:

Def.7.3 Let A be a set and $\succsim_D$ a binary relation on A. If $\succsim$ is defined as in Def.7.2 then the structure $\langle A, \succsim_D \rangle$ is a *partly connected order* iff, for all x, y, z $\in$ A, the following axioms hold:

1 $(x \succsim_D y \;\&\; y \succsim_D z) \rightarrow x \succsim z$
2 $(x \succsim y \;\&\; y \succsim z) \rightarrow x \succsim z$.

The first of these axioms is the counterpart of Ax.2 of Def.5.4 and is reasonably clear. It means that when each of the pairs x, y and y, z can be compared directly and the results are as indicated then all possible indirect comparisons of x with z yield results in accordance with $x \succsim z$. This axiom is sufficient to show that $\succsim$ is connected (see proof below). The second axiom merely states that $\succsim$ is transitive. This is to some extent an admission of failure in the light of the statement made above about wanting the axioms to be as simple as possible. I have not found a simpler alternative. As we show very easily below, this definition is technically quite satisfactory, but whether or not it is possible to simplify it further is an interesting minor problem.

7.2.3 Properties of a Partly Connected Order

We now state and prove two theorems.

Th.7.1 If the structure $\langle A, \succsim_D \rangle$ is a partly connected order and $\succsim$ is defined as in Def.7.2 then the structure $\langle A, \succsim \rangle$ is a weak order.

Proof: Connectedness.

Assume:

$\neg\ x \succsim y.$

For this to occur the antecedent in Def.7.2 must hold for some w, i.e.:

$(\exists a)(a \succsim_D x \ \&\ y \succsim_D a).$

Then, from Ax.1 we have: $y \succsim x$.

Transitivity. Axiom 2. ∎

Th.7.2 If the structure $\langle A,\succsim_D\rangle$ is a partly connected order and $\succsim$ is defined as in Def.7.2 then for all $x,y \in A$ the following holds:

$x \succsim_D y \rightarrow x \succsim y.$

Proof:

Assume:

$x \succsim_D y \ \&\ \neg\ x \succsim y.$

Then:

$x \succsim_D y \ \&\ (\exists a)(a \succsim_D x \ \&\ y \succsim_D a).$

Hence, using Ax.1:

$x \succsim a \succ a \succsim y.$

With Ax.2 this gives $x \succsim y$. ∎

This establishes the prime results we need. We have two relations $\succsim_D$ and $\succsim$. The first is interpreted in terms of a direct mutual interaction of two objects. Because of this it cannot obtain between all pairs of objects that we would wish to describe, in normal parlance, as standing in the relation of "being at least as heavy as" (or the corresponding relation for some other measurable property), and so cannot represent that relation. The second is interpreted in terms of an indirect comparison of two objects via their behaviour with all other objects with which both can be caused to interact. It does connect all pairs of objects (in the intended set) and so is a plausible candidate to represent the appropriate quantitative relation. In summary, we see $x \succsim y$, not $x \succsim_D y$, as carrying the meaning of "x is at least as heavy as y". As the second theorem shows a direct operational test can supply sufficient evidence that one object is at least as heavy as another, but it does not by itself supply the meaning of that assertion.

We now turn to dealing with the other problem that was introduced in Section 6.1, the problem of constructibility.

7.3 Concatenation and Constructibility

7.3.1 The Method of Replicas

There is a commonly employed method of avoiding the problem of constructibility. It is indicated in the wording of Ax.4 of Def.7.1 and involves informal appeal to the notion of replicas. It is usual to assume (though with no explicit provision for this in the axioms) that for any element x there is another element x', itself a member of A, that is a replica of x. The term $x \circ x$ is then taken to represent the concatenation of object x with its replica, that is it is $x \circ x$ identified with another object $x \circ x'$. According to the usual meaning of "replica", a replica of an object is distinct from it and so $x \circ x'$ is constructible. It will be a physical object having just the properties we intuitively expect. It will be of twice the magnitude of the original object. A similar provision comes from identifying objects such as (x o y) o x with (x o y) o x', (x o y) o (x o z) with (x o y) o (x' o z), and so on. The idea seems to be that in the interpretation of the structure the term x o y is to be understood along the lines:

I.7.2 x o y is an object resulting from the concatenation of object x either with object y, or with an object y_r, where y_r is an object obtained from y by replacing any constituent of y that is common to x by a replica of that constituent that is itself distinct from x, from any constiuent of x, and from any other constituent of y.

However there are some difficulties associated with this solution. For one thing it has serious consequences for the membership of A. It is quickly seen that for any object x in A must have not merely one but an infinite number of replicas - x', x'',x'''.... say. x' is needed to allow the identification of x o x with x o x', x'' to allow the identification of (x o x') o x or (x o x') o x' with (x o x') o x'', and so on without limit. This is to say the least an implausibly strong requirement. Quite apart from its implications for the population of the universe, it makes it impossible to establish an extensive measure on a finite set. This is plainly refuted by what occurs in practice.

A second, related, difficulty is that the existence of an infinite number of replicas of an object x gives an infinite number of composite objects x o x', x o x'', x o x'''....etc. any one of which may be chosen to be identified with x o x. It seems to be quite arbitrary which is taken to be x o x. But different choices give different operations. For instance the operation o_1 for which, say, among other correspondences we have:

$$x\ o_1\ x = x\ o_1\ x',$$

is a different operationfrom o_2 for which:

$$x\ o_2\ x = x\ o_2\ x''$$

is the appropriate identification. Hence there is no unique operation on A corresponding with concatenation. Associated with the pair $\langle A, \succsim\rangle$ there is an infinite number of structures $\langle A, \succsim, o_1\rangle$, $\langle A, \succsim, o_2\rangle$, $\langle A, \succsim, o_3\rangle$ and so on, all of which support the same homomorphism into $\langle Re, \geq, +\rangle$. Selection of a particular one as the structure to represent the quantitative attribute in question would be arbitrary.

We could try to avoid these particular difficulties by arguing that there are unique entities to which x o x, (x o y) o x and so on correspond, but that these entities are abstract, not physical, objects. This would remove the need for arbitrary choices of objects that are required to stand in as physical manifestations of these terms and the concatenation operation would be uniquely specified. For example, suppose that there are three wholly distinct physical objects a, b and c among the members of A. Concatenation will produce four more physical objects, namely a o b, b o c, a o c and a o b o c (let us assume for the present point that o is associative and commutative), but in addition will give rise to the abstract objects, a o a, a o a o a, a o a o a o a, ..., a o a o b, a o a o a o b, ... and so on. By virtue of the closure of o there still is an infinite number of these objects, but they take up no space in the physical world. The move reduces the overcrowding and restores the possibility of defining a structure on the basis of a finite set of physical objects.

However apart from this, the solution has little to commend it. The idea that the possibility of ordering a set of physical objects for the purposes of extensive measurement is dependent upon the existence of an abstract superstructure of the kind we have described is most implausible. And the resulting mixture in the membership of A is not at all appealing. The set now contains not only physical objects but abstract objects of various kinds, some like (x o x) formed by concatenation of physical objects with themselves, some like (x o x) o (y o y) formed by combining other abstract objects, some like (x o x) o y formed by combining abstract objects with physical objects and so on. This ontological heterogeneity leads to difficulties of comparability considerably more serious than those already disposed of in the earlier sections of this chapter. There are now members of A ontologically unsuited to appearing in an operational test at all. We cannot place the object that x o x is supposed to represent on the pan of a balance, and so it cannot be compared, even indirectly with any other member of A. The methods of Section 7.2 are of no avail in this situation. If we were to adopt this solution we would need some means for determining the truth value of an expression such as $x \circ x \succsim y$. We would need, that is, an interpretation of the general expression $x \succsim y$ that is broad enough to deal with several disparate situations. It would have to cover cases where both x and y are physical objects, cases where both are abstract objects, and cases where there is one of each. This would extend the meaning of $\succeq$ so far beyond our usual understanding of the comparative terms of physical measurement as to be unacceptable. Expressions like "a is as heavy as or heavier than b" have meaning only when both a and b are physical objects.

7.3.2 Structures with Restricted Concatenation

A much more promising line of attack on this problem is to accept that the concept of concatenation simply does not apply for a pair of objects that are not wholly physically distinct, and to rule out as candidates for membership of A entities such as x o x and (x o y) o x and the rest, whether viewed as surrogate physical objects, or as abstract objects, or in any other way. This implies that, contrary to Def.7.1, o is not a closed operation on A, that is, it is no longer a function from (A x a) to A. Instead it is a function from a certain proper subset, B say, of (A x a) to A, where B is the set of physically independent pairs of members of A. x o y is constructible and is accepted as a member of A only if the pair (x, y) is a member of B. This is so if, and only if, the objects represented by x and y are wholly distinct. Pairs such as (x, x), (x o y, x) and so on are not members of B and so x o x, (x o y) o x, and so on are excluded from A. It is necessary, therefore, to modify the definition of the measurement structure to take account of these constraints on the membership of A.

Examples of this approach are to be found in the literature. One is to be found in Luce and Marley (1969). These authors have developed a theory to take account of limited concatenation. They give a set of axioms for an extensive structure based on a quadruple $\langle A, B, \succsim, o\rangle$ where B is a subset of AxA. However, it is clear that their theory is designed to cope with a different problem from the one that concerns us here. The axioms of their system do not in general exclude pairs like (x, x) and (x, x o y) from membership of B and so troublesome objects such as x o x, and x o (x o y) are still admitted to A. Luce and Marley's interest is in systems where restrictions on concatenation are due to experimental limitations. In an actual situation there will be, for example, a limit to the total weight of bodies that can be placed on the pan of a balance without disrupting its functioning, or a limit to space available for joining rods end to end, and so on. It is difficulties of this sort that their theory is designed to deal with and they are different from the ones that presently interest us. They are due to contingent features of the particular systems involved and and not to something intrinsic to the idea of concatenation of physical objects. Because of this B must be incorporated in their structure as an extra primitive term. A broadly similar scheme, devised to meet the same problem, is to be found in Narens (1986). He gives a definition of a *positive concatenation structure* that includes the following axiom of *local definability* (Sec. 7):

If $x \circ y$ is defined and $x \geq w$ & $y \geq z$ holds then $w \circ z$ is defined.

This axiom does not deal with the problem of constructibility. For suppose that x and y are wholly distinct both from each other and from w and z, but that w and z have some part in common (or are identical). Then the antecedent may hold even though $w \circ z$ is nonconstructible.

In our case the restrictions on the membership of A are an inherent feature of the concatenation of physical objects, and we seek to express them more economically in terms of the concatenation function itself. We now consider how this can be done.

Consider the following as an initial suggestion.

Def.7.4 Let A be a set and o a binary operation partially defined on A. The structure $\langle A, \circ \rangle$ is a *proper concatenation structure* iff for all x, y $\in$ A, $x \circ y \in A$ unless one of the following holds:

1 $x = y$
2 $(\exists u)(x = y \circ u \ \vee \ y = x \circ u)$
3 $(\exists u)(\exists v)(\exists w)(x = u \circ v \ \& \ y = u \circ w)$

This is intended to capture all possible cases of pairs that cannot be concatenated, namely those for which (1) the two members of the pair are the same object, (2) one is part of the other, and (3) the two have some part in common. Unfortunately, as we show below, it happens that the success of this definition depends on conditions that are not available to us in Def.7.1. These conditions are that the concatenation operation should be commutative and associative, that is that the following should hold for all x, y, z:

(i) $x \circ y = y \circ x$, and

(ii) $x \circ (y \circ z) = (x \circ y) \circ z$.

Neither (i) nor (ii) is guaranteed in Def.7.1. For example Ax.2 of Def.7.1 expresses the equivalence of $x \circ (y \circ z)$ and $(x \circ y) \circ z$ but not their identity. This is sometimes referred to as weak associativity (relative to $\succsim$). If (ii) was added to the theory, Ax.2 would become redundant, since it would then be covered by the reflexivity condition stated in Ax.1. Similar remarks apply to the commutativity of o. Weak commutativity, i.e. the condition $x \circ y \sim y \circ x$, is implied by Def.7.1 though it is not needed as a separate axiom since it follows from Axioms 1, 2 and 4. (A proof can be found in Krantz *et al* (1971), p.78.) But (strong) commutativity, condition (i), cannot be derived.

In the absence of these two properties Def.7.4 is inadequate. For example the two bodies represented by $x \circ y$ and $y \circ x$, or the two represented by $x \circ (y \circ z)$ and $(x \circ y) \circ z$, are made up of exactly the same components, and so cannot be concatenated. But in the absence of commutativity and associativity the inference that they are identical is blocked, and so clause (§) is powerless to exclude $(x \circ y) \circ (y \circ x)$ and $(x \circ (y \circ z)) \circ ((x \circ y) \circ z)$ and other such objects from membership of A. We might have been tempted to suppose that this can be remedied by adding further clauses to Def.7.4 to cover the extra cases. We might try for example to expand clause 1 in the following way:

1′ either $(x = y)$
or $(\exists u)(\exists v)(x = u \circ v \;\&\; y = v \circ u)$
or $(\exists u)(\exists v)(\exists w)(x = u \circ (v \circ w) \;\&\; y = (u \circ v) \circ w)$

But it is quickly seen that this does not advance us very far. It turns out that this augmented definition is not sufficient to capture pairs such as $(x \circ y) \circ z$ and $(y \circ x) \circ z$, or $(x \circ (y \circ z)) \circ w$ and $((x \circ y) \circ z) \circ w$ which are further examples of pairs of objects with the same constituents. Adding yet more clauses to deal with these is of course fruitless since the problem recurs for yet more complex objects. Corresponding problems occur with clauses 2 and 3 of Def.7.4. If (i) and (ii) do not hold; they too are inadequate to cover all the intended cases and there is no apparent way of amending them so that they do.

In a later section of this chapter we shall be discussing the general significance of associativity and commutativity and I shall argue there that there are no compelling reasons to prevent us from adopting these conditions as axioms of an extensive structure. In that case Def.7.4 would be sufficient. However Miller, in a private communication, has pointed out that it is possible to deal with the present problem in the absence of these conditions by means of a recursive definition, and I am indebted to him for the following suggestion.

Def.7.5 Let A be a set and o a binary function from a subset B of $A \times A$. The structure $\langle A, o\rangle$ is a *proper concatenation structure* iff the following is satisfied. For all $x, y, z \in A$:

1 $(x, x) \notin B$
2 $(x, y) \notin B \rightarrow (y, x) \notin B$
3a $(x, y) \notin B \rightarrow (x, y \circ z) \notin B$
3b $(x, y) \notin B \rightarrow (x, z \circ y) \notin B$
3 $(x, y) \in B$ unless excluded by conditions 1, 2, 3a or 3b.

Equipped with Def.7.5 we are now in a position to define an extensive structure that appears to be free of the problems we have been discussing in the last three sections. The following is a possible definition.

Def.7.6 Let A be a nonempty set, $\succsim_D$ a binary relation on A and o a binary operation partially defined on A. Then the triple $\langle A, \succsim, o\rangle$ is a *positive extensive structure* iff the following axioms are satisfied for all $x, y, z \in A$:

1 $\langle A, o\rangle$ is a proper concatenation structure.
2 $\langle A, \succsim_D\rangle$ is a partially connected order.
3 $x \circ (y \circ z) \sim (x \circ y) \circ z$.
4a $x \succsim y$ iff $x \circ z \succsim y \circ z$.
4b $x \succsim y$ iff $z \circ x \succsim z \circ y$.
5 $x \circ y \succ x$.
6 $(\exists e)\{(x \sim e) \vee (\exists u)(x \sim u \circ e)\}$.

Axs.1, 2 incorporate the results of the analyses of Sections 6.2 and 6.3. Axs.3, 4, 5 are unchanged from Def.7.1. Ax.6 replaces the Archimedean axiom and gives a necessary and sufficient condition for the structure to be a well determined order.

Def.7.6 retains the weak forms of associativity and commutativity. However the discussion leading up to it did show that there may be some technical advantage in having the strong forms of these properties. I now wish to investigate what significance there might be in the difference between the two sets of properties.

7.4 Associativity and Commutativity of the Concatenation Operation

We now examine the implications of adopting the strong conditions:

(i) $x \circ y = y \circ x$, and (ii) $(x \circ y) \circ z = x \circ (y \circ z)$

in addition to:

(iii) $x \circ y \sim y \circ x$, and (iv) $(x \circ y) \circ z \sim x \circ (y \circ z)$.

The decision is whether or not the pair $x \circ y$ and $y \circ x$, or the pair $x \circ (y \circ z)$ and $(x \circ y) \circ z$, may be counted as representing the same object. At first glance this may look like a question of metaphysics, a question of whether or not we can assume the existence of some fixed entity beneath changes in the configuration of the component parts. But I do not think that it is a fundamental issue of this sort. It is simply a technical decision about what to take as members of the set A. If we adopt the strong conditions, (i) and (ii), members of A are single objects and <u>sets of configurations</u> of objects. If the weak conditions only, (iii) and (iv), are adopted, then members of A are single objects and <u>individual</u> configurations of objects (or possibly smaller sets of configurations). The decision can be made on pragmatic grounds according to what we want the theory for. Such decisions are commonplace in science. In physics for example we might treat a sample of gas in a container as a fixed object for the purposes of determining its mass, despite the fact that the configuration of its component molecules is changing continuously through random motion, whereas, for the purposes of statistical mechanics, we treat individual configurations as distinguishable entities. In this section I want to examine the question in relation to a range of typical applications of measurement structures. I shall argue that for the purposes of measurement theory there are no good reasons for maintaining a distinction between $x \circ y$ and $y \circ x$, or between $x \circ (y \circ z)$ and $(x \circ y) \circ z$, and that no essential characteristic of measurement is lost from the theory by adopting (i) and (ii).

In order to decide whether or not it is correct to equate terms like $x \circ y$ and $y \circ x$, and terms like $(x \circ y) \circ z$ and $x \circ (y \circ z)$, we must first be clear how the differences between them are to be understood. It is difficult to generalize from discussions in the literature about the nature of the distinctions. It is usual to offer specific examples, illustrating how each pair of terms may be interpreted in particular cases. In some cases the interpretation is obvious and plausible enough, in others perhaps a little contrived, but there is little attempt to spell out general principles. It may well be worthwhile, therefore, to begin by surveying a range of examples to find out if there are any general features that bear upon the decision about (i) and (ii). We shall begin with the question of commutativity.

7.4.1 Concatenation and commutativity

We must first note the difference between what I call *sequential* concatenation and *non-sequential* concatenation. It is clear that for the distinction between $x \circ y$ and $y \circ x$ to be significant the concatenation procedure must be serial in character. It must produce a sequential arrangement of some kind. Most commonly this is a one-dimensional array of the objects involved. Consider for example the concatenation of rods in the measurement of length. Two rods, represented by x and y, can be abutted in a linear order along a specified axis so that x o y corresponds with the arrangement:

Fig.7.1

A third rod z can be added to give:

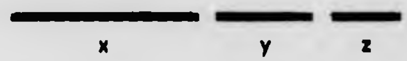


Fig.7.2

which is represented by (x o y) o z, and so on. However if the order of the objects x and y is reversed we obtain two different configurations as shown here.

Fig.7.3

Fig.7.4

The formalism conveniently allows us to record the distinction between these latter arrangements and those in Fig.7.1 and Fig.7.2. We can represent them differently, that in Fig.7.3 by y o x and that in Fig.7.4 by (y o x) o z.

Another illustration is provided by a series arrangement of resistors in an electrical circuit, which gives a one-dimensional array.

(It is not necessary for the resistors to be set along a straight line, of course; the array need not be confined to one dimension of physical space.) We can thus distinguish between the two configurations shown here:

Fig.7.5

Fig.7.6

representing them, again, by x o y and y o x, or between:

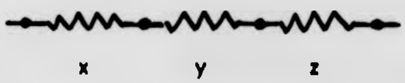


Fig.7.7

Fig.7.8

representing them by (x o y) o z and (y o x) o z. These are clear examples of sequential concatenation.

In other examples however, the matter is not so clear. Consider the concatenation of weights. This is achieved by placing them together on the pan of a balance. What is the distinction between x o y and y o x in this case? As far as spatial configuration is concerned there are not merely two, but indefinitely many ways in which two bodies x and y may be arranged in the pan. This concatenation procedure does not generate a serial arrangement. Now there may be reasons for wanting to distinguish the different configurations, but the point is that it cannot be done in the formal language. The one-dimensional pattern of the symbolism does not match the two-dimensional pattern of arrangements on the scale pan. In such a case it is difficult to see how the pair of terms x o y and y o x can be assigned unambiguous yet distinct meanings.

We might suggest an alternative method of concatenation. In this, bodies are arranged vertically in a column in the pan. The resulting configuration does have a serial character and we may for example refer to the following two cases:

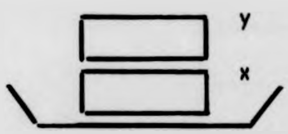


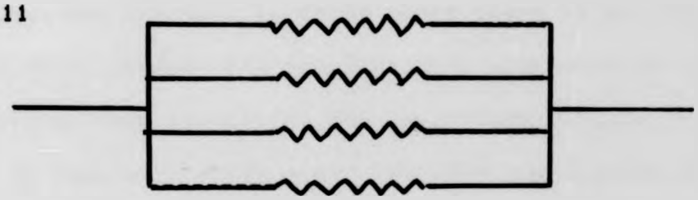
Fig.7.9

Fig.7.10

as x o y and y o x respectively. That these two arrangements are equivalent in their effect on the balance is a significant result. That the downward thrust on the pan is independent of the order in which the bodies are piled up is part of the classical theory of statics, depending in part on Newton's Third Law about action and reaction. In view of this it may be thought useful to distinguish x o y and y o x. Nevertheless, the suggested concatenation method hardly corresponds with usual practice in using a balance; it is little more than a contrivance to provide the formalism with a job.

Other examples of non-sequential concatenation abound. Consider the construction of a scale of electrical conductance, as opposed to a scale of resistance. For this the appropriate concatenation operation involves placing resistors - perhaps more judiciously referred to in this context as conductors - in parallel, rather than in series. However, parallel arrangements do not in general determine a sequence of any kind. The conventional way in which they are often depicted in circuit diagrams, i.e. as a set of resistors lying side by side in order in a plane as in Fig.7.11, might suggest otherwise.

Fig.7.11



But this is merely an accident of two-dimensional representation. The set of resistors depicted in this figure may simply be connected as a bundle between A and B. Further examples are the concatenation of laminar objects in fundamental measurement of area and the concatenation of solid objects in fundamental measurement of volume. There are many more. In none of these examples is any sequence of objects necessarily produced and so there is no clear distinction in use between the terms $x \circ y$ and $y \circ x$.

An alternative reading of the difference between the terms $x \circ y$ and $y \circ x$ is sometimes proposed, with a view to capturing all examples of concatenation in the sequential category. On this reading the order represented by the formalism is temporal, not spatial. The term $x \circ y$ reflects the sequence of operations in the assembly of the object it represents. For example it can stand for the object that results from first placing object x on the pan of a balance, and then adding object y, in which case $y \circ x$ stands for the result of performing these operations in the reverse order. Then we may if we so choose treat the end results of the different sequences of operations as different objects, i.e. we may stipulate that $x \circ y \neq y \circ x$.

This idea applies well in cases where there is an explicit temporal element. Take as an example the concatenation operation with pendulums in the measurement of time intervals. The usual idea is that $x \circ y$ stands for the sequence of events in which first the pendulum x performs one swing, and then the second y, starting its swing as the first finishes. In this case $x \circ y$ and $y \circ x$ clearly represent different sequences of events. However as a general principle it seems very weak. There are many cases where it makes little sense. When two rods are joined end to end the union is brought about at a single instant. There is no sense in which one is added later than the other. The same applies to joining resistors in a circuit. Even where there may be a definite temporal sequence, as with the case of weights added to a scale pan, it may not be an essential feature of the operations involved. We can in principle conduct weighing operations successfully on the basis of concatenating objects by placing them in the pan simultaneously.

Thus it can be argued that the facility for representing different configurations by maintaining a formal distinction between x o y and y o x is not of general use throughout the whole range of possible applications of extensive structures. At best the distinction is appropriate only for cases of sequential concatenation, since as far as the non-sequential kind there simply is no point to it. However, even for the case of sequential concatenation the possibility of maintaining the distinction in question cannot be taken for granted. There is a further complication that seems not to have been generally noticed. I shall discuss it in the next section.

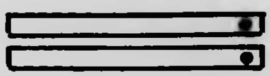
7.4.2 Commutativity and Reversibility

In those cases where concatenation produces a spatial sequence, the commutativity of the operation, whether weak or strong, is dependent upon the existence of what is perhaps a more basic property of the objects concerned which we will refer to as *reversibility*. The point can be explained very simply with reference once more to the example of rigid rods and length measurement.

In that particular example we took x o y and y o x to represent different arrangements of rods x and y along an axis. Now in constructing a scale of length it is not usual to specify a particular axis. Rather we assume that space is isotropic with respect to the relevant properties of rigid rods. We assume that the length of a rigid rod is unaffected by its orientation in space. A common view is that this is a matter of convention since it is untestable. Suppose we wish to test whether or not a given rod does change when rotated through some angle, say 90 degrees. We compare it with some other rod in each of the two positions and find that the two are equivalent in both orientations. But the difficulty is obvious. The second rod has itself to be rotated through the same angle in order to make the comparisons, and the same question must be answered with respect to the effect of the rotation on it. The same applies if a third rod is employed, and so on. Provided that the same equivalences are observed in both orientations, the question is undecidable by operational tests and we may conventionally adopt the view that no change has taken place.

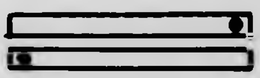
However this does not exhaust the matter. The case is different for rotations of 180 degrees. A straightforward empirical test is available for this. Suppose two rods x and y are found to be equivalent in length when compared as in the following diagram.

Fig.7.12



It is observed that they are still equivalent when one is inverted with respect to the other as now shown:

Fig.7.13



It is logically possible that this might not have been so. There are examples of measurable attributes that have directional characteristics and for which the results are different. One is magnetic dipole moment. Imagine a system for comparing bar magnets, in which two magnets are placed equidistant from a magnetometer aligned as indicated in Fig.7.14.

Fig.7.14

If the magnets are of equal moment, i.e. if they are equivalent, the magnetometer shows null deflection. The system is in effect a magnetic balance. Now if one of the magnets is reversed to give the arrangement:

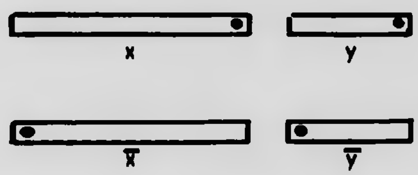
Fig.7.15

the reading of the magnetometer changes. The two magnets are no longer observed to be equivalent.

We can represent the situation symbolically in this way. We denote the object that results from reversing object x by $\bar{x}$. Clearly, $x = \bar{\bar{x}}$ holds for all objects. But the question that needs to be decided for any given application is: what is the relation between x and $\bar{x}$? Recalling the outcome of the tests we have just been describing we note that the relation:

(i) $x \sim \bar{x}$

holds in the case of rigid rods, whereas in the case of the magnets it does not. We can now show very simply that (i) implies weak commutativity. First note, by inspection of the following figure, that:



$x \circ y$

$\bar{x} \circ \bar{y} = \overline{y \circ x}$

(ii) $\bar{x} \circ \bar{y} = \overline{y \circ x}$

Then from (i) we have:

(iii) $x \sim \bar{x}$ & $y \sim \bar{y}$.

This with Ax.3 Def.7.1 gives:

(iv) $x \circ y \sim \bar{x} \circ y$ & $\bar{x} \circ y \sim \bar{x} \circ \bar{y}$.

This with Ax.1, Def.7.1 (transitivity) gives:

(v) $x \circ y \sim \bar{x} \circ \bar{y}$

This with (ii) gives:

(vi) $x \circ y \sim \overline{y \circ x}$

From (i) we have:

(vii) $\overline{y \circ x} \sim y \circ x$

Finally, from (vi) and (vii) and transitivity, we have:

(viii) $x \circ y \sim y \circ x$

Hence (i), which we may call the condition of *weak reversibility* implies weak commutativity.

We might choose to regard x and $\bar{x}$ as the same object, i.e. to assume *strong* reversibility:

(ix) $x = \bar{x}$

It is trivial to show that we then get (strong) commutativity; (ii) and (ix) give:

$$x \circ y = \overline{x \circ y} = \bar{y} \circ \bar{x} = y \circ x.$$

That is, (strong) reversibility implies that (strong) commutativity. Thus, decisions on commutativity in the structure may well be preempted by, usually implicit, assumptions on reversibility.

Reversibility on the other hand is independent of commutativity. The reverse implications do not hold. Nonreversibility is compatible with commutativity, weak or strong. Refer again to the example of bar magnets. Of the following arrangements the first two are equivalent to each other in their magnetic effects, whereas the third is equivalent to neither.

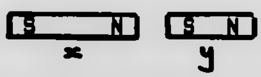


Fig.7.16

Fig.7.17

Fig.7.18

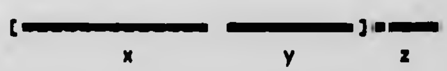
Here, $x \circ y \sim y \circ x$ holds but not $x \circ y \sim \bar{x} \circ \bar{y}$.

Thus we see that the overall picture with regard to commutativity is far from uniform from one application to another. In some cases the attribution of significance to differences in the order of concatenation is simply a contrivance. In some others, where the differences are genuinely significant, it seems that the standard theory is insufficiently equipped to deal with them because related features, like reversibility, are not represented in the language. It is hard to escape the conclusion that the facility for making the distinction between x o y and y o x is not essential for the formal representation of extensive measurement, but rather it is an additional feature that is exploited, somewhat haphazardly, when the opportunity arises.

7.4.3 Concatenation and Associativity

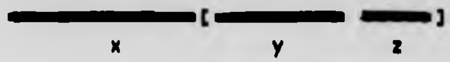
Broadly similar considerations apply to the matter of associativity though overall there seems rather less to say. As with commutativity it is difficult to discern any general principles operating throughout the whole range of applications. Again some sequential pattern associated with the concatenation process appears to be called for but, in some applications at least, the feature that provides the sequence to which commutativity applies, may not be the appropriate one for associativity. Consider again the normal method of concatenating rods, where the distinction related to commutativity was made in spatial terms. In contrast there is no distinction to be made between (x o y) o z and x o (y o z), that is between:

Fig.7.19



and:

Fig.7.20



in terms of spatial configuration. In this case the only distinction worth considering appears to be the temporal order in which the complete object is assembled. In Fig.7.19 y was first joined to x, and then z was added. In Fig.7.20 y and z were first joined and then x added. We have the option as far as the formalism is concerned of treating the results as distinct objects. However, given that they are identical in physical constitution and spatial arrangement, and are equal in length, it is not clear what motivation there is for insisting on the distinction. As for an example like the concatenation of weights in a scale pan, there appears to be no scope at all for a distinction between (x o y) o z and x o (y o z) either in spatial or in temporal terms.

It is worth recalling that in Ellis's system (described in Section 3.2.1), where concatenation of rods involves joining them at right angles, changing the order of concatenation does make a significant difference to the shape of the resulting composite object. Figs. 3.2 and 3.3 show the difference in shape between (x o y) o z and x o (y o z) and in the whole range of possible configurations shown in Figs. 3.2 to 3.7 we see the combined effects of associativity and commutativity. Nevertheless even here it is not entirely clear what the point may be of claiming that a pair of terms such as $(x \circ y) \circ z$ and $x \circ (y \circ z)$ refer to different objects. For as we pointed out in the discussion, a single term like $(x \circ y) \circ z$ itself refers to a whole set of different shapes. (In the case of $(x \circ y) \circ z$, it will be recalled a range of shapes can be generated by rotating z around the major chord of $x \circ y$ - see Fig.3.2.) To be prepared to treat any two of these as the same object while insisting that any two represented by $(x \circ y) \circ z$ and $x \circ (y \circ z)$ respectively are different seems arbitrary.

I conclude that although distinctions between $x \circ y$ and $y \circ x$ and between $(x \circ y) \circ z$ and $x \circ (y \circ z)$ are significant in many structures, the fact that they are not so in all of them indicates that the features to which they refer are separate from those that characterize extensive measurement in general. That is to say, nothing that is essential to understanding the foundations of measurement is lost by adopting the condition of strong commutativity and associativity.

7.5 An Extensive Measurement Structure with Boolean Operations

It will be clear that, given the strong conditions of associativity and commutativity, what we have called a proper concatenation structure has strong affinities with a Boolean algebra. The set A of a proper concatenation structure can be partitioned into two subsets. The first, A_o say, consists of the atoms, i.e. those elementary members of A that are not composed of other members. The second, A_1, the complement of A_o, consists of the molecules, those that are composites (of members of A_o).

For example:

(i) If $A_o = \{a, b\}$ then $A_1 = \{a \circ b\}$
(ii) If $A_o = \{a, b, c\}$ then $A_1 = \{a \circ b, a \circ c, b \circ c, a \circ b \circ c\}$
etc.

In general A_o may be defined in terms of A and o by:

Def.7.7 For all $x \in A$, $x \in A_o$ iff $\neg(\exists y)(\exists z)(x = y \circ z)$.

Set A can now be represented by $P(A_o)$, the power set of A_o. There would be a one-one correspondence between A and $P(A_o)$ if there was a member of A to correspond with the empty set. We have not mentioned this possibility so far but in most if not all cases of extensive measurement it is usual to assume the existence of an "empty" object that can appear in operational tests and to which the value zero is assigned on an additive scale. In operations with a balance for example it is posible to perform tests in which a single object is placed in one pan while the other is left empty. The observed behaviour of a normal balance under these cicumstances is consistent with treating the empty pan as an extra member of A_o and assigning it its own place in the order. Since every other object tips the balance against the empty pan, this position is fixed unambiguously at the appropriate extremity of the order. The admission of the empty pan to the set need not be seen as adding a special object different in ontological status from the other members. We can avoid difficulties of this sort by regarding the members of the ordered set as something like pan-states, rather than as the objects themselves, and an empty pan will correspond with a normal member of such a set.

It is clear that the fact that the empty pan can be regarded in this way is an empirical discovery. Outcomes other than those actually observed with the empty pan are logically possible. It could have been that only bodies above a certain position in the order tipped the balance against the empty pan while with the rest it was the reverse. We would then no doubt be ready to postulate the existence of two sorts of weight, "positive" and "negative". The empty pan would be assigned to the position in the order that separated one sort from the other. More bizarre results might have been obtained. Suppose it turned out that not all bodies that tipped the balance against the empty pan were higher in the order than all which failed to. This result would not disturb the original order, but it would be inconsistent with treating the empty pan as an additional object to be included in it. Thus whether or not there is a physically realizable object that can play the part of the empty object in an operational test is open to investigation from one example to another. For length, for example, the limiting case is a straight rod whose ends coincide; which is, of course, no rod at all. For time interval, it is a pair of events that are simultaneous. For resistance it is a short circuit, while for conductance it is an open circuit. Both of these are in the last analysis ideal concepts - a true open circuit, for instance, requires there to be a perfect vacuum between the circuit terminals - but they can in general be realized sufficiently closely in practice.

If the set of objects is augmented in this way we may reformulate Def.7.6 in Boolean terms as follows:

Def.7.8 Suppose that A is a nonempty set, that $\succsim_D$ is a binary relation on P(A), the power set of A, and that $\succsim$ is defined as in Def.7.2. The structure $\langle A, \succsim_D \rangle$ is an extensive structure iff for every X, Y, Z $\in$ P(A):

1 $\langle P(A), \succsim_D \rangle$ is a partially connected order.
2 If $X \cap Z = Y \cap Z$ then $X \succsim Y$ iff $X \cup Z \succsim Y \cup Z$.
3 If $X \neq Y$ then $X \cup Y \succ X$

4 $(\exists Z)(X \neq \emptyset \;\&\; (X)(\exists Y)(X \sim Y \cup Z)$

In this scheme A is a set of physically independent objects. (It corresponds with A_o in the earlier notation but there is no reason to retain the subscript.) The relations $\succsim_D$ and $\succsim$ now connect not pairs of members of A but pairs of subsets of A including of course the singletons. Thus expressions such as $x \succsim y$ and $x \circ y \succsim z$ of the theory of Def.7.6 become $\{x\} \succsim \{y\}$ and $\{x, y\} \succsim \{z\}$ in this formulation.

7.6 Cancellation axioms

Although the positive extensive structure of Def.7.8 appears to provide a solution to the problems arising from restrictions on physical concatenation that we have discussed in this chapter, the outcome is not totally satisfactory. There is another intriguing difficulty. Consider the following example, devised by Kraft, Pratt and Seidenberg (1959).

Suppose we have a subset $A_1 = \{a, b, c, d, e\}$ of A_0 and that operational tests yield the following results:

(i)
$$\begin{aligned} a &\succsim_D bc \\ bd &\succsim_D ac \\ ce &\succsim_D ab \end{aligned}$$

where we use an obviously abbreviated notation to represent specific subsets of A_1. If an additive representation n exists we will have:

$$\begin{aligned} n(a) &\geq n(b) + n(c) \\ n(b) + n(d) &\geq n(a) + n(c) \\ n(c) + n(e) &\geq n(a) + n(b) \end{aligned}$$

Then, by addition and cancellation, we will have:

$$n(d) + n(e) \geq n(a) + n(b) + n(c)$$

This in turn implies that a test will give the result:

(ii) $$de \succsim_D abc$$

We can obtain this result from Def.7.6; but, what is perhaps surprising, we need the structural axiom, Ax.4, to do so. It turns out that (ii) is not a consequence of the conjunction of (i) and Axioms 1, 2 and 3 alone. It is consistent with it but so also its its negation:

(iii) $$\neg(de \succsim_D abc), \quad (\text{i.e. } abc \succ_D de).$$

The trouble lies with the cancellation axiom of Def.7.6:

(Def.7.6, Ax.3) $x \circ (y \circ z) \sim (x \circ y) \circ z$.

Attempts to use this axiom to rule out (iii) are blocked by the intrusion of undefined terms like aab. This is one more manifestation of the problem of constructibility. Yet clearly if (i) and (iii) hold together there can be no additive representation.

The original discussion of the problem by Kraft *et al* related to probability representations, and subsequent comment appears to have been confined to that context, with the attendant suggestion that its existence indicates something peculiar to probability (see e.g. Fine, 1973, p.25). However in principle the problem could arise for any extensive property. The reason why it has not attracted attention in the area of physical measurement is, I suspect, that it is not usual to consider the ordering of finite arbitrarily structured sets. It is standard for theories to include structural axioms, such as our Ax.3, that impose strong conditions on the membership of the set A. In general terms they require that, for some given member or members of A, there exist others related in such a way that there will be a sufficiently large set of operational test results available to determine numerical inequalities in favour of an additive representation. It is worth showing briefly how this can work.

As we have already remarked in an earlier section of the chapter a typical requirement is that for any object, x, in the set, there are other objects, x', x''...say, in the set that are replicas of it. This would provide more than enough to solve the present problem. As it happens augmenting A_1 by the addition of a single replica would be sufficient to avert the unfortunate result in (iii). Suppose we add a replica of a to give the larger set A_2 = {a, a', b, c, d, e, }. With the addition of a ~ a' to (i) the proof of (ii) is simple.

Proof.

We have from (i) with repeated use of Ax.2,:

$$bde \succsim_D ace \succsim_D a'ce \succsim_D a'ab,$$

which by transitivity gives:

$$bde \succsim_D a'ab.$$

From this, with further appeal to Ax.2 and (i), we obtain:

$$de \succsim_D a'a \succsim_D a'bc \succsim_D abc.$$

The required result:

$$de \succsim_D abc$$

follows by transitivity. ∎

Thus the problem is avoided in sufficiently structured sets. Kraft *et al* in their 1959 paper, in discussing a structure for representing probability, opt for a solution based on the addition of what they refer to as "irrelevant events", such as the tossing of coins, to the original set of events on which a measure is to be provided (p.418). Nevertheless it is rather unsatisfactory to have to rely on structural axioms for this. It is surely desirable to have a set of nonstructural axioms that are sufficiently strong to encapsulate the essential relational properties underlying a measurable attribute. Suppose that the original five objects in set A_1 in our example were weights, and that the results in (i) were obtained with a balance. If result (iii) was also obtained then it would be clear that something was wrong, without any need for recourse to tests employing a sixth object. We would reject the conjunction of (i) and (iii) as inconsistent with our intuitions about weight irrespective of the contingent matter of whether or not replicas were available. It is reasonable to require as much of an adequate formal theory.

There is no point in trying to deal with the problem simply by adding a further cancellation axiom:

(iv) $x \succsim_D yz \;\;\&\;\; yv \succsim_D xz \;\;\&\;\; zw \succsim_D xy \;\rightarrow\; vw \succsim_D xyz$

For we may reinstate the problem by devising an example involving a larger number of elements. Suppose we have a set {a, b, c, d, e, f, g, h, j, k} and that the following results are obtained:

(v)
$$\begin{aligned} ab &\succsim_D cde \\ bcg &\succsim_D ade \\ cdh &\succsim_D abe \\ dej &\succsim_D abc \\ aek &\succsim_D bcd \end{aligned}$$

If an additive representation exists, we should expect the result:

(vi) $ghjk \succsim_D abcde$

But (vi) cannot be derived from (v) and the nonstructural axioms of Def.7.6 even when the latter are augmented by (iv). Scott and Suppes (1958) have shown that it is not possible to solve the problem with a finite number of (first order) cancellation axioms.

Scott (1964, Theorem 4.1) gives a set of axioms that are free of structural requirements and that are sufficient to establish a measure on a finite Boolean algebra. However as the author points out, his theory is not itself expressible solely in Boolean terms. It includes an operation for the *algebraic sum* of (the characteristic functions of) two sets. This is equivalent to their union only in the case of disjoint sets. It is of some interest to investigate whether or not it is possible to solve the problem within a Boolean framework. In the next, final section I give what I believe to be an adequate theory that meets this condition. The section is largely taken up with the construction of mathematical apparatus that appears to be needed. The resulting theory is rather cumbersome and may well not have much intuitive appeal. However in view of the importance of this problem it seems worthwhile to show that it is possible to find a solution along these lines in the almost certain hope that there is scope for simplification.

7.7 An extensive structure with cancellation

A central feature of the theory to be described here, in keeping with the fact that it involves higher order techniques, is that separate pieces of information given in a sequence such as in (i) or (v) of the last section can be handled together as a package. We shall continue with (i) as an illustrative example. The sequence of tests - the comparison of a with bc, of bd with ac, and so on - of which the expressions in (i) report the outcome is regarded as a single experiment. We denote it by E. The expression in (ii) also reports the outcome of a test - the comparison of de with abc - and we count that too as an, albeit shorter, experiment. It is in some sense a derivative of E. Very roughly we can describe it as being obtained from E by adding up and then cancelling out as appropriate the various occurrences of each element in the sequence of tests in E, and using the residue to construct another test. The theory described below gives a formal procedure for doing this. For any experiment E there is a unique experiment R(E), referred to as the *reduction* of E, derivable by means of the reducing function R that is defined in the theory. In addition the theory includes an axiom that places constraints on the joint outcome of the pair E and R(E), and these rule out the possibility of disastrous results of the sort that the conjunction of (i) and (iii) of the last section would represent.

The definition of R(E) is lengthy and technically rather elaborate, and rather than giving a full formal statement immediately it is probably better to introduce it in stages with some explanatory comment. Something shorter and less cumbersome is almost certainly possible, but, as we have already remarked, the chief interest here is in the demonstration that a satisfactory rule can be given in Boolean terms, rather than in the mathematical details.

Some preliminaries are necessary and first we define some terms.

7.7.1 <u>Preliminary Definitions</u>

a) An ordered pair $t = (X, Y)$, where $X, Y \in P(A)$, is a *test*. If $X \succsim Y$ holds then t is a *positive* test (relative to $\succsim$). Otherwise t is *negative*. [The axioms in the complete definition will require that $\emptyset \succsim \emptyset$ holds and so in particular the empty test $t_\emptyset = (\emptyset, \emptyset)$ is positive.]

b) A finite sequence of tests $E = \langle t_1, t_2, t_3, \ldots, t_n \rangle$ is an *experiment*. If, and only if, all t_k in E are positive, E is a *positive* experiment. If, and only if, all t_k in E are negative, E is a *negative* experiment.

c) An experiment may consist of a single test, $\langle (X, Y) \rangle$, in which case it is a *simple* experiment. However we extend the notion of a simple experiment to include cases such as:

$$\langle (X, Y), (\emptyset, \emptyset), (\emptyset, \emptyset), \ldots, (\emptyset, \emptyset) \rangle,$$

or even:

$$\langle (X, Y), (X, Y), \ldots, (X, Y), (\emptyset, \emptyset), (\emptyset, \emptyset), \ldots, (\emptyset, \emptyset) \rangle$$

that is, sequences for which $t_k = t_1$ or $t_\emptyset$ for all k.

As an example of the way in which these terms are used consider our original problem. We can re-express (i) by asserting that the sequence:

(vii) $E_1 = \langle(a, bc), (bd, ac), (ce, ab)\rangle$

is a positive experiment. According to the definition to be given below the reduction of E_1 is the experiment:

(viii) $R(E_1) = \langle(de, abc), (\emptyset, \emptyset), (\emptyset, \emptyset)\rangle$.

This is a simple experiment. To anticipate the final details, the relevant axiom of the theory will impose the condition that if E_1 is positive and $R(E_1)$ is simple, then $R(E_1)$ is also positive. This would imply the result we are after, namely de $\succeq$ abc.

It is reasonably clear in informal terms what the reducing function R must achieve in its action on the general experiment E:

(ix) $E = \langle(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\rangle$.

First, the elements that occur throughout the sets $X_1, X_2, \ldots, X_n$, must be sorted, keeping stock of the number of occurrences of each one, and the same for those occurring throughout $Y_1, Y_2, \ldots, Y_n$. Second, equal numbers of occurrences of elements common to both totals must be cancelled. To these ends we define two functions.

(a) The cancelling function

One is the cancelling function, which is straightforward. Given the experiment E from above we define another experiment C(E) by:

Def.7.9

$$C(E) = \langle (X_1-[X_1 \cap Y_1],\ Y_1-[X_1 \cap Y_1]),\ (X_2-[X_2 \cap Y_2],\ Y_2-[X_2 \cap Y_2]), \ldots\ldots, (X_n-[X_n \cap Y_n],\ Y_n-[X_n \cap Y_n])\rangle$$

This simply cancels common elements from each side of individual tests. For the case where $X_k \cap Y_k = \emptyset$ for all k clearly C(E) = E. We shall refer to C(E) as the *cancellation* of E.

(b) The sorting function

The second, the sorting function, is more intricate. The idea is to start with the two sequences:

$$X_1,\ X_2,\ X_3,\ldots,X_n \qquad \text{and} \qquad Y_1,\ Y_2,\ Y_3,\ldots,Y_n,$$

whose terms come from the tests that make up an experiment, and derive from them, by an inductive process, a pair of corresponding sequences:

$$X_1^\cap,\ X_2^\cap,\ X_3^\cap,\ldots,X_n^\cap \qquad \text{and} \qquad Y_1^\cap,\ Y_2^\cap,\ Y_3^\cap,\ldots,Y_n^\cap.$$

The second pair of sequences are related to the first in the following way. $X_1^\cap$ is made up of elements that are members of at least one of $X_1, X_2,\ldots,X_n$, $X_2^\cap$ is made up of only those elements that are members of at least two of $X_1, X_2,\ldots,X_n$, $X_3^\cap$ at least three, and so on. Some term $X_k^\cap$, and then all higher terms, in the new sequence may may be identical with Ø. Indeed this will be so unless there is at least one element that is a member of every X_i of the original sequence. More generally, if there are exactly k occurrences of some element x distributed among X_1, X_2, ...X_n, then x is a member of sets X^n_1, X^n_2.....X^n_k but of no others in the sequence. An entirely similar description applies to the Y-sequences. On the basis of the new sequences we define the sorting of E, S(E) by:

(x) $S(E) = \langle(X^n_1, Y^n_1), (X^n_2, Y^n_2), (X^n_3, Y^n_3), ...(X^n_n, Y^n_n)\rangle$.

We can return once again to our example to illustrate this. The action of the sorting function on:

$E_1 = \langle(a, bc), (bd, ac), (ce, ab)\rangle$

leads to:

(xi) $S(E_1) = \langle(abcde, abc), (\emptyset, abc), (\emptyset, \emptyset)\rangle$

And before proceeding with the description of the inductive procedure for deriving S(E) from E, it is worth continuing a little further with the example to show how the cancelling and sorting functions are to be used together to achieve the reduction of an experiment. If we now use the cancelling function on $S(E_1)$ we obtain:

(xii) $CS(E_1) = \langle(de, \emptyset), (\emptyset, abc), (\emptyset, \emptyset)\rangle$, Then one more sorting produces:

(xiii) $SCS(E_1) = \langle(de, abc), (\emptyset, \emptyset), (\emptyset, \emptyset)\rangle$.

This is $R(E_1)$, the reduction of E_1 introduced in (viii). The action of this trio of operations, in the order shown, is perfectly general. The first sorting arranges the elements in such a way that the subsequent cancelling operation has maximum impact. After the cancellation, there is no element to be found anywhere among the members of the left hand terms of the tests that is also a member of a right hand term, and vice versa. Following the cancellation there is in general, scope for further sorting. However, once this is done, no further application of either C or S will produce anything new. That is to say, for any experiment E:

(xiv) $CSCS(E) = SSCS(E) = SCS(E)$.

In anticipation of a correct formal definition of the sorting function we explicitly define the reduction of E by:

Def.7.10 If E is an experiment, the *reduction* of E is the experiment $R(E) = SCS(E)$.

As was already noted above in the case of E_1 the reduction $R(E_1)$ is a simple experiment. This does not happen in general for arbitrary E.

We now deal with the definition of the sorting function.

The task is to define a procedure for replacing any n-termed sequence of sets by its corresponding n-termed sorted sequence, i.e

$$F = \langle X_1,\ X_2,\ X_3 \ldots\ldots X_n \rangle \quad ==> \quad F^*(n) = \langle X_1^n,\ X_2^n,\ X_3^n \ldots\ldots X_n^n \rangle$$

This is achieved by a procedure in which, effectively, F is processed one term at a time. It generates a list of sequences $F^*(1)$, $F^*(2)$.... in which the first, then the first two, then the first three, etc. terms of F are sorted. This is indicated in this array of (all n-termed) sequences.

$$\begin{array}{lll}
F(1) = \langle X_1, \emptyset, \emptyset, \ldots\ldots\ldots\ldots \emptyset \rangle & ==> & F^*(1) = \langle X_1^1,\ \emptyset,\ \emptyset, \ldots\ldots\ldots\ldots \emptyset \rangle \\
F(2) = \langle X_1, X_2, \emptyset, \ldots\ldots\ldots\ldots \emptyset \rangle & ==> & F^*(2) = \langle X_1^2,\ X_2^2,\ \emptyset, \ldots\ldots\ldots\ldots \emptyset \rangle \\
\quad\quad \ldots\ldots\ldots & & \quad\quad \ldots\ldots\ldots \\
\quad\quad \ldots\ldots\ldots & & \quad\quad \ldots\ldots\ldots \\
F(m) = \langle X_1, X_2, \ldots X_m, \emptyset, \ldots\ldots \emptyset \rangle & ==> & F^*(m) = \langle X_1^m,\ X_2^m, \ldots\ldots X_m^m,\ \emptyset \ldots\ldots \emptyset \rangle \\
F(m+1) = \langle X_1, X_2, \ldots X_m, X_{m+1}, \emptyset \ldots \emptyset \rangle & ==> & F^*(m+1) = \langle X_1^{m+1}, X_2^{m+1}, \ldots\ldots X_m^{m+1}, X_{m+1}^{m+1} \ldots \emptyset \rangle \\
\quad\quad \ldots\ldots\ldots & & \quad\quad \ldots\ldots\ldots \\
\quad\quad \ldots\ldots\ldots & & \quad\quad \ldots\ldots\ldots \\
F(n) = \langle X_1, X_2, \ldots\ldots\ldots\ldots\ldots X_n \rangle & ==> & F^*(n) = \langle X_1^n,\ X_2^n, \ldots\ldots\ldots\ldots\ldots X_n^n \rangle
\end{array}$$

Each sequence on the right is the sorted version of the corresponding sequence on the left. It is easy to see what happens near the beginning.

1 In $F^*(1)$, $X_1^1 = X_1$ and so trivially $F^*(1) = F(1)$.

2 In $F^*(2)$, $X_1^2 = X_1 \cup X_2$ and $X_2^2 = X_1 \cap X_2$. Any element that occurs twice (i.e. in both X_1 and X_2) appears in both X_1^2 and X_2^2. That is, X_2^2 stores the overspill from X_1^2.

3 In $F^*(3)$, any element that occurs three times (i.e. in X_1, X_2 and X_3) appears in X_1^3, X_2^3 and X_3^3.

And so on. For the definition to work inductively we need a rule for generating $F^*(m+1)$ from $F^*(m)$. To put it informally, $F^*(m+1)$ is obtained by bringing in the next set, X_{m+1}, from the original sequence F and distributing its members appropriately along F^*m. Any member of X_{m+1} not already represented in X_1^m can be accommodated there, but any other is passed to the right along $F^*(m)$ until it reaches a position where it can be accepted, namely the first position not already occupied by one of its own kind. Hence, the first term of $F^*(m+1)$, X_1^{m+1}, comes from adding to X_1^m any member of X_{m+1} it does not already have, and thus it is given by:

(xv) $$X_1^{m+1} = X_1^m \cup X_{m+1}$$

There will be a residue if $X_1^m \cap X_{m+1}$ is nonempty. The second term of the new sequence, X_2^{m+1}, will be the result of the union of X_2^m with this residue, i.e:

(xvi) $$X_2^{m+1} = X_1^m \cup (X_1^m \cap X_{m+1}).$$

In general, for k > 1, the membership of the kth term of the (m+1)th sequence, X_k^{m+1}, will be the made up of the membership of the kth term of the mth sequence, X_k^m, plus any new members from X_{m+1} that are already in the (k-1)th term of the mth sequence. That is:

(xvi) $X_k^{m+1} = X_k^m \cup (X_{k-1}^m \cap X_{m+1})$.

This can all be summarized in the following recursive definition of the general term X_k^m.

Def.7.11

(1) $X_1^1 = X_1$

$X_1^m = X_1^{m-1} \cup X_m$ for $1 < m \leq n$

(2) $X_k^1 = \emptyset$

$X_k^m = X_k^{m-1} \cup (X_{k-1}^{m-1} \cap X_m)$ for $1 < m \leq n$

$\Big\}\ 1 < k \leq n$

We are now in a position to define the sorting function.

Def.7.12 If E is the experiment:

$$E = \langle (X_1, Y_1), (X_2, Y_2), \ldots\ldots.(X_n, Y_n)\rangle,$$

the sorting of E is the experiment:

$$S(E) = \langle (X_1^n, Y_1^n), (X_2^n, Y_2^n), \ldots\ldots.(X_n^n, Y_n^n)\rangle$$

where the terms X_k^n are as given in Def.7.11, and similarly for Y_k^n.

This completes the preliminaries for the definition of an extensive structure. We now state the axioms of the structure.

7.7.2 Definition of a Boolean Structure with Cancellation

Def.7.2 is repeated here for reference. This, it will be recalled, is a definition of an ordering relation $\succsim$ in terms of a direct operational relation $\succsim_D$ which meets the problem of noncomparability.

Def.7.2 for all $X, Y, W \in P(A)$

$$X \succsim Y \equiv (W)((W \succsim_D X \ \& \ Y \succsim_D W) \rightarrow (X \succsim_D W \ \& \ W \succsim_D Y))$$

On the basis of this we define an extensive structure as follows:

Def.7.13 Suppose that A is a nonempty set, that $\succsim_D$ is a binary relation on P(A), the power set of A, and that $\succsim$ is defined as in Def.7.2. The structure $\langle A, \succsim_D \rangle$ is an extensive structure iff for every $X \in P(A)$ and for any experiment E on P(A):

1 $X \succsim \emptyset$.
2a If E is positive and R(E) is simple then R(E) is positive.
2b If E is negative then R(E) is not positive.

We state and prove two basic theorems.

Th.7.4 If the structure $\langle A, \succsim_D \rangle$ is an extensive structure and $\succsim$ is defined as in Def.7.2 then the structure $\langle A, \succsim \rangle$ is a weak order.

The proof is given over the page.

Proof of Th.7.4.

We make use of notation abbreviated in an obvious way as indicated in the following figures. (Here variables x, y, etc. stand for mutually disjoint members of P(A), i.e. for sets, and not for members of A itself, as they have done so far.)

Fig.7.21

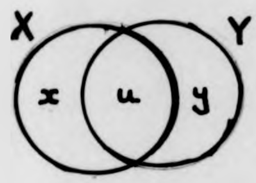
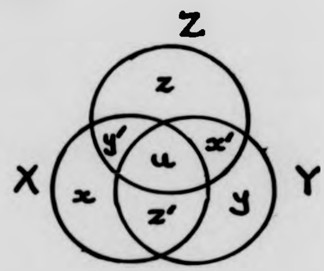


Fig.7.22



Connectedness.

Suppose that for some X, Y that neither $X \succsim Y$ nor $Y \succsim X$ holds. Then:

$$E = \langle (X, Y), (Y, X) \rangle$$

is negative. Using the notation of Fig.7.21, put X = xu and Y = yu so that:

$$E = \langle (xu, yu), (yu, xu) \rangle.$$

Then:

$$S(E) = \langle (xyu, xyu), (u, u) \rangle,$$

which gives:

$$CS(E) = \langle (\emptyset, \emptyset), (\emptyset, \emptyset) \rangle$$

which in this instance is also R(E). By Ax.1 this is positive, which contradicts Ax.2b.

Transitivity.

Suppose that for some X, Y, Z we have:

$X \succsim Y$ and $Y \succsim Z$.

Then:

$E = \langle (X, Y), (Y, Z) \rangle$

is positive. Using the notation of Fig.7.22, put $X = xy'z'u$, etc. giving:

$E = \langle (xy'z'u, x'yz'u), (x'yz'u, x'y'zu) \rangle$.

Then:

$C(E) = \langle (xy', x'y), (yz', y'z) \rangle$,
$SC(E) = \langle (xyy'z', x'yy'z), (\emptyset, \emptyset) \rangle$, and
$CSC(E) = R(E) = \langle (xz', x'z), (\emptyset, \emptyset) \rangle$.

This is simple, and hence, by Ax.2 is positive. Now consider the experiment:

$E' = \langle (X, Z) \rangle = \langle (xy'z'u, x'y'zu) \rangle$.

The reduction of this is:

$R(E') = \langle (xz', x'z) \rangle$,

Comparison with R(E), which has just been shown to be positive, shows that R(E') is positive. Hence, by Ax.2b, E' is not negative, and so:

$X \succsim Z$. ∎

Th.7.5 For all $X, Y \in P(A)$ $X \cup Y \succsim X$.

Proof: Consider the experiment:

$E = \langle (X \cup Y, X) \rangle = \langle (xyu, xu) \rangle$ (in the notation of Fig.7.21).

The reduction of this is:

$R(E) = \langle (y, \emptyset) \rangle$

which by Ax.1 is positive. Hence E must be positive which gives:

$X \cup Y \succsim X$. ∎

We round this off by stating and proving a theorem equivalent to the cancellation axiom, Ax.4 of Def.7.6.

Th.7.6 For all $X, Y, Z \in P(A)$ if $X \cap Z = Y \cap Z$ then $X \succeq Y$ iff $X \cup Z \succeq Y \cup Z$.

Proof: The notation of the following figure will be used.

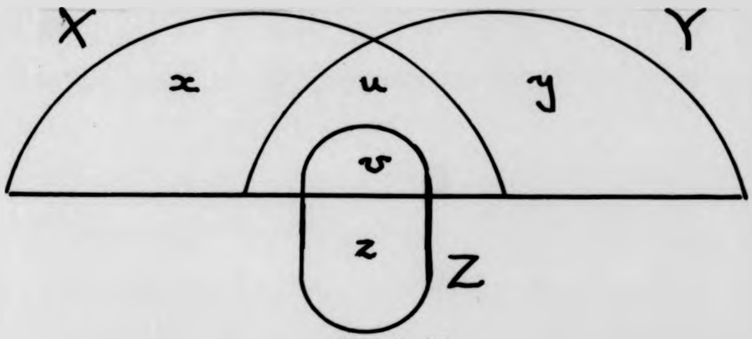


Fig.7.23

Fig.7.23 takes account of the hypothesis $X \cap Z = Y \cap Z$. The two experiments:

$E = \langle (X, Y) \rangle = \langle (xuv, yuv) \rangle$, and:

$E' = \langle (X \cup Z, Y \cup Z) \rangle = \langle (xuvz, yuvz) \rangle$

both have the same reduction $\langle (x, y) \rangle$. Hence, by Ax.2, if one is positive the other cannot be negative and the required result follows. ∎

This completes the account of the definition and basic properties of the Boolean structure. We have seen that it appears to deal satisfactorily with the problem of cancellation.

I end with a brief suggestion of a possible further development of this theory. At the end of Section 2.1.3 it was pointed out that it is possible to establish a well determined order on the basis of an operational *equivalence* relation; that is, an operational *ordering* relation is not a necessary requirement for extensive measurement. If this is so it should be possible to give an adequate formulation of an extensive structure as an alternative to Def.7.13 in terms of $\sim$ rather than $\succsim$, and it would be of some interest to do so. Initial indications are that this can be achieved, but the details have yet to be worked out.

Bibliography

Barford, N. C. [1985]: *Experimental Measurements: Precision, Error, and Truth*. Chichester: John Wiley and Sons Ltd.

Bridgman, P. [1931]: *Dimensional Analysis.* Newhaven: Yale University Press.

Campbell, N. R. [1920]: *Physics: the Elements.* Cambridge: Cambridge University Press. (Republished as *Foundations of Science*, New York: Dover Publications, Inc., 1957.)

Carnap, R. [1966]: *Philosophical Foundations of Physics.* New York: Basic Books, Inc.

Coombs, C. H. [1952]: 'A Theory of Psychological Scaling', *Bull.Engng.Res.*, 34, Ann Arbor: University of Michigan Press.

Edwards, D. F. A. [1971]: *Electronic Measurement Techniques.* London: Butterworths.

Ellis, B. [1966]: *Basic Concepts of Measurement*. London: Cambridge University Press.

Erlich, P. [1981]: 'The Concept of Temperature and its Dependence on the Laws of Thermodynamics', *Am. J. Phys.*, 49, pp. 622-632.

Field, H. H. [1980]: *Science Without Numbers*. Oxford: Basil Blackwell.

Fine, T. [1973]: *Theories of Proability*. New York: Academic Press.

Giles, R. [1964]: *Mathematical Foundations of Thermodynamics.* New York: Macmillan.

Goodman, N. [1951]: *The Structure of Appearance.* Harvard University Press.

Goodman, N. [1972]: 'Order from Indifference', in *Problems and Projects*, Indianapolis and New York: Bobbs Merrill Company, Inc.

Holman, E. W. [1969]: 'Strong and Weak Extensive Measurement', *J. Math. Psychol.*, 6, pp. 286-293.

Holman, E. W. [1973]: 'Extensive Measurement without an Order Relation', *J. Phil.*, 41, pp. 361-373.

Home, D. [1977]: 'Concept of Temperature without the Zeroth Law', *Am. J. Phys.* 45, pp. 1203-1205.

Kraft, C. H., Pratt, J. W. and Seidenberg, A. [1959]: 'Intuitive Probability on Infinite Sets', *Ann. Math. Stat.*, 30, pp. 408-419.

Krantz, D. H. [1967]: 'Extensive Measurement in Semiorders', *Phil. Sci.*, 34, pp. 348-362.

Krantz, D. H., Luce, R. D., Suppes, P. and Tversky, A. [1971]: *Foundations of Measurement, Vol. I*. N.Y. and London: Academic Press.

Kyburg Jnr., H. E. [1984]: *Theory and Measurement*. London: Cambridge University Press.

Luce. R. D. [1956]: 'Semiorders and a Theory of Utility Discrimination', *Econometrica*, 24, pp. 178-91.

Luce, R. D. and Marley, A. A. J. [1969]: 'Extensive Measurement when Concatenation is Restricted and Maximal Elements May Exist', in S. Morgenbesser, P. Suppes and M. G. White (*eds.*) *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*. New York: St. Martin's Press.

Luckhardt, S. C. and Kessler, J. O. [1971]: 'Equivalence of the Zeroth and Second Laws of Thermodynamics', *Am. J. Phys.*, 39, pp. 1496-1498.

Narens, L. [1973]: 'Measurement without Archimedean Axioms', *J. Phil.*, 41, pp. 374-393.

Narens, L. [1986]: *Abstract Measurement Theory*. M.I.T. Press.

Scott, D. [1964]: 'Measurement Structures and Linear Inequalities', *J. Math. Psychol.*, 1, pp. 233-247.

Scott, D. and Suppes, P. [1958]: 'Foundational Aspects of Theories of Measurement', *J. Symbolic Logic*, 23, pp. 113-128.

Stevens, S. S. [1946]: 'On the Theory of Scales of Measurement', *Science*, 103, pp. 677-680.

Suppes, P. [1957]: *Introduction to Logic.* Princeton: D. Van Nostrand Company Inc.

Zemansky, M. W. and Dittman, R.H. [1981]: *Heat and Thermodynamics*, New York: McGraw-Hill. (5th edition of Zemansky, M. W. [1951]: *Heat and Thermodynamics*, New York: McGraw-Hill.)